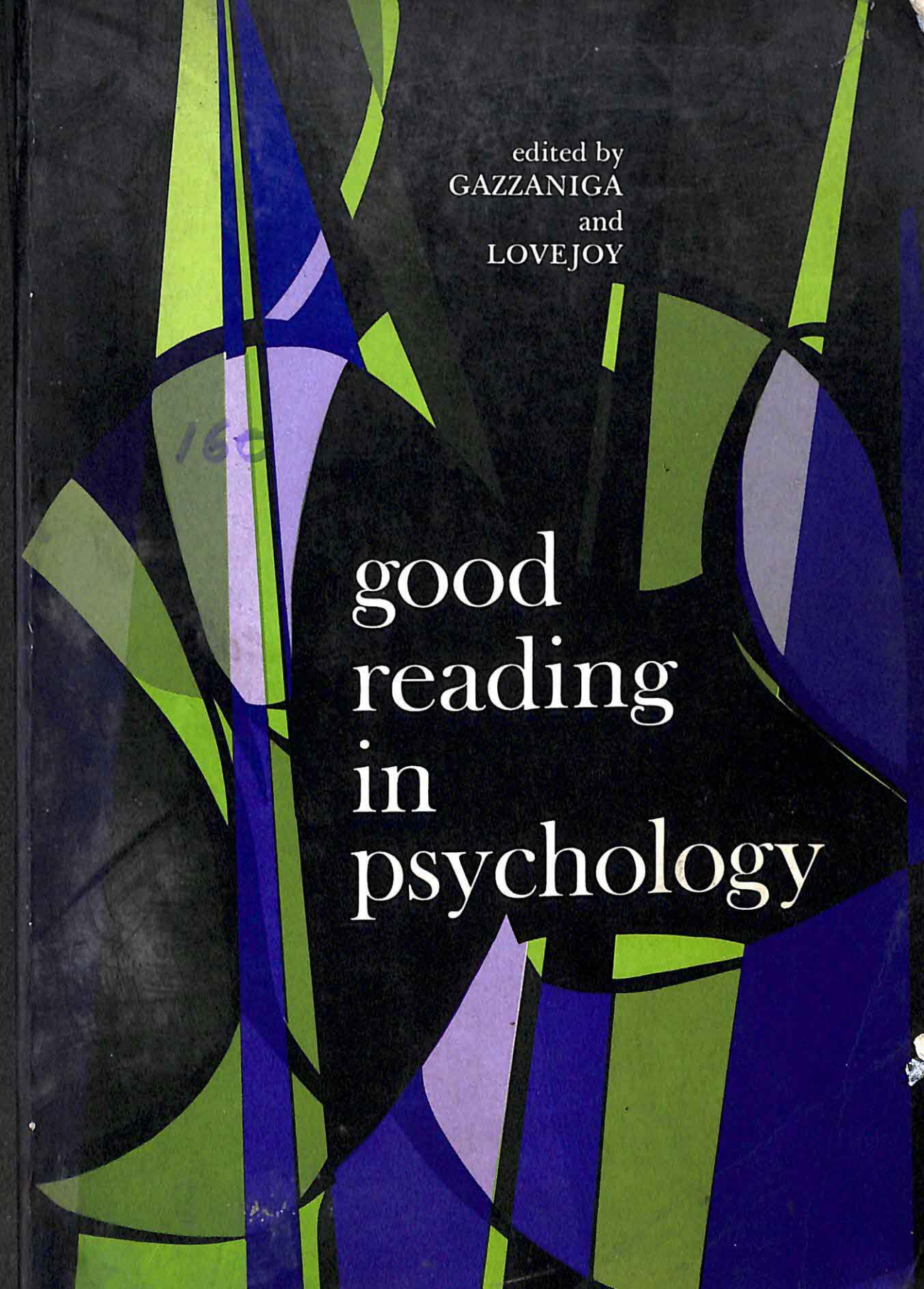




edited by
GAZZANIGA
and
LOVEJOY

160

good
reading
in
psychology

233
15.3.72

160

good

reading

in

psychology

PRENTICE-HALL SERIES IN EXPERIMENTAL PSYCHOLOGY

James J. Jenkins, Editor

✓
good
reading
in
psychology

edited by

Michael S. Gazzaniga

New York University

and

Elijah P. Lovejoy

University of California, Santa Barbara



PRENTICE-HALL, INC., *Englewood Cliffs, New Jersey*

S.C.E.R.T., West Bengal

Date 15.3.72

Acc. No. 2336

© 1971 by Prentice-Hall, Inc., Englewood Cliffs, N.J.

All rights reserved. No part of this book may be reproduced in any form or by any means without permission in writing from the publisher.

13-360503-5

Library of Congress Catalog Card Number: 77-127060

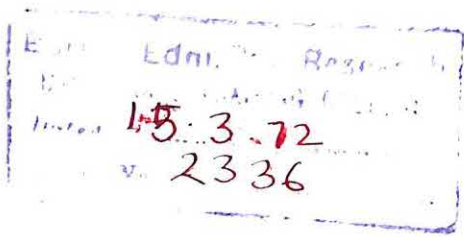
Printed in the United States of America

Current printing (last number):

10 9 8 7 6 5 4 3 2 1

PRENTICE-HALL INTERNATIONAL, INC., London
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., Sydney
PRENTICE-HALL OF CANADA, LTD., Toronto
PRENTICE-HALL OF INDIA PRIVATE LIMITED, New Delhi
PRENTICE-HALL OF JAPAN, INC., Tokyo

150
G+2



This collection has been assembled with the following five premises in mind.

1. In order to enjoy psychology one should find the very best that the field has to offer. This involves searching out the greatest thinkers and writers, and then selecting the best of the works they have produced.
2. Students are more likely to read these papers if they are assembled together into one binding rather than scattered in a library.
3. Most existing "readers" have tried to give a representative idea of the entire field of psychology. This often results in a feast with an enormous number of small courses, each of which relates only weakly to the others in the meal. Such a repast is very difficult to digest.
4. Experimental psychology produces incredible numbers of reports of carefully conducted little experiments, jam-packed with elegant data. It has produced considerably fewer pieces of large-scale integration which pull together a variety of results from many programs of research. Wide-ranging, integrative essays should have an important place in any psychological library.
5. It is impossible to cover the whole field of psychology, even at a very shallow level, in a single book of readings. What is possible however, and what we have tried to do, is to single out a small number of especially exciting questions and areas, and to present several papers in each one.

If we have left out your favorite integrative essay, please let us know about it. The next edition will be so much the better for your suggestions.

Good reading!

Acknowledgments

This collection depends in the most obvious way on the permissions so graciously granted by the authors of the works reprinted and by their publishers. We are very grateful for these permissions. Putting together a collection of this sort requires a surprisingly large amount of detail work and organization, and we would probably have foundered had it not been for the extensive and expert assistance of Melanie McLees and Donald Burns. We are most grateful to them for giving us the benefit of their very great skills in this endeavor.

S.C.E R T., West Bengal

Date

Acc. No.....

CONTENTS

PART ONE

nature vs nurture 1

1. *ON PSYCHOSEXUAL DEVELOPMENT* W. Simon / J. Gagnon, 5
2. *ARE THERE SUCH THINGS AS MENTAL DISEASES?*
J. B. Watson, 26
3. *CRIME AS DESTINY* J. Lange, 30
4. *ARE THERE ANY HUMAN INSTINCTS?* J. B. Watson, 40
5. *IDENTICAL TWINS: MENTAL TRAITS* A. Scheinfeld, 53
6. *MONKEYS REARED IN ISOLATION WITH PICTURES AS VISUAL
INPUT: EVIDENCE FOR AN INNATE RELEASING MECHANISM*
J. B. Sackett, 63
- ✓ 7. *PHYSIOLOGICAL AND CULTURAL DETERMINANTS OF
BEHAVIOR* N. E. Miller, 70

PART TWO

free will and determinism 87

8. *A PHILOSOPHER LOOKS AT SCIENCE* J. G. Kemeny, 90
9. *MIND, BRAIN, AND HUMANIST VALUES* R. W. Sperry, 106
10. *FREEDOM OF ACTION IN A MECHANISTIC UNIVERSE*
D. M. MacKay, 121

PART THREE

structure of knowledge 139

- ✓ 11. *COMPUTING MACHINERY AND INTELLIGENCE*
A. M. Turing, 142

12. *THE BEE-HUNTERS OF HULHORST* N. Tinbergen, 165
13. *WHAT IS THINKING?* J. B. Watson, 176
14. *THE PROBLEM OF SERIAL ORDER IN BEHAVIOR*
K. S. Lashley, 180
15. *THE MISBEHAVIOR OF ORGANISMS* K. Breland / M. Breland, 199
16. *QUALITATIVE VERSUS DIRECTIONAL CUES IN TWO FORMS
OF DIFFERENTIATION* C. Dobrzecka / G. Szwejkowska /
J. Konorski, 206
17. *THE UNIT OF ANALYSIS* G. A. Miller / E. Galanter / K. Pribram, 211
18. *COGNITIVE MAPS IN RATS AND MEN* E. C. Tolman, 221

PART FOUR

biological determinants of behavior 245

19. *REGULATIVE FACTORS IN THE ORDERLY GROWTH OF
RETINO-TECTAL CONNEXIONS* R. Sperry / E. Hibbard, 248
20. *RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S
STRIATE CORTEX* D. Hubel / T. Wiesel, 259
21. *RELATIONS BETWEEN THE CENTRAL NERVOUS SYSTEM AND
THE PERIPHERAL ORGANS* E. von Holst, 277
22. *SELF-STIMULATION OF THE BRAIN* J. Olds, 285
23. *THE RIDDLE OF FRONTAL LOBE FUNCTION IN MAN*
H. Teuber, 307

PART FIVE

control of human behavior 339

24. *SLIPS OF THE TONGUE* S. Freud, 342
25. *CHANGING THE GAME FROM "GET THE TEACHER" TO
"LEARN"* R. Hamblin / D. Buckholdt / D. Bushell / D. Ellis /
D. Ferritor, 347
26. *SELF-CONTROL PROCEDURES IN PERSONAL BEHAVIOR
PROBLEMS* I. Goldiamond, 363
27. *SELF-CONTROL* B. F. Skinner, 379

PART SIX

frontiers in psychology 391

- 28. *THE GENERAL PROPERTIES OF LANGUAGE* N. Chomsky, 393
- ✓ 29. *REINFORCEMENT THEORY* D. Premack, 412
- ✓ 30. *ON WAR AND PEACE IN ANIMALS AND MAN* N. Tinbergen, 426
- ✓ 31. *HYPNOTISM* W. James, 444
- 32. *PAIN-INDUCED FIGHTING IN THE SQUIRREL MONKEY*
N. Azrin / R. Hutchinson / D. Hake, 459
- ✓ 33. *THE OPPORTUNITY FOR AGGRESSION AS AN OPERANT
REINFORCER DURING AVERSIVE STIMULATION*
N. Azrin / R. Hutchinson / R. McLaughlin, 461
- 34. *LEARNING OF VISCERAL AND GLANDULAR RESPONSES*
N. E. Miller, 473
- ✓ 35. *CLEVER HANS: A CASE STUDY OF SCIENTIFIC METHOD*
R. Rosenthal, 498
- 36. *SIMILARITIES OF INHIBITION IN THE DIFFERENT SENSE
ORGANS* G. von Bekesy, 520
- 37. *MEMORY-SCANNING: MENTAL PROCESSES REVEALED BY
REACTION-TIME EXPERIMENTS* S. Sternberg, 538

good

reading

in

psychology

Well Mrs. Jekyll, if this condition has never appeared in your family before, perhaps you've let your son watch too many horror shows. I suggest that you get him interested in a more healthy area... like medicine.



PART ONE

nature vs. nurture

Psychology may be thought of as the series of endeavors which result from man's unending preoccupation with himself. Fascinated by his own essential nature, he has been engaged ever since the start of recorded history in elaborate theorizing about the very nature of man and the forces which affect him. One expression of this interest is a question which recurs in various forms across all of psychology and indeed throughout life. Am I as I am because of my genetic makeup—because of a detailed set of instructions written into my chromosomes at conception, instructions which are now being followed and are subject to no significant change from without? Or am I rather a man, like all other men, and different from them solely because of my differing experience in this world? To what extent is my personality, my character and my very self the result of innate genetically determined factors, and to what

extent am I the result of the environment in which I have grown up and the experiences which I have had? This dichotomy, often referred to as the nature-nurture question or the heredity vs environment question, has long had a fascination for men interested in themselves—and that means for all men.

In the articles that follow both extreme views are presented, along with some arguments and evidence supporting them. At the same time the important twin-method is illustrated. This method takes advantage of the fact that occasionally nature produces a pair of men with identical genetic makeup—identical twins, also called monozygotic twins. (Monozygotic simply means that both individuals in a pair of identical twins originated in a single cell, formed at conception. This cell split into two identical cells each of which eventually matured into a complete being—one of the identical twins.) This happy circumstance means that we can look analytically at the extent to which the future behavior of these two individuals is similar even when they have grown up under very different conditions. To the extent that such similarity is found, and to the extent that there is more similarity for identical twins than for nonidentical twins, we can conclude that the genetic influences are significant and are strong enough to overcome differences in environment. Two of the articles below, that by Lange and that by Scheinfeld, used the twin technique.

The first article in this section, "On Psychosexual Development," by Simon and Gagnon, looks at human sexual behavior, which is often thought of as expressing violently powerful innate forces. Disagreeing with the Freudian emphasis on innate patterns of sexual motivation, the authors argue forcefully that we *learn* many of those behaviors. As you read, try to analyze the role of *learning* in your own past history.

The second selection is taken from Watson's book *Behaviorism* which had an enormous effect, both positive and negative, upon the development of psychology. Watson's ideas are very instructive not only because he writes well but especially because he has deliberately taken a polar position: the position that hereditary constraints are negligible and that all important changes in human development are due to environmental forces. Watson's book had a profound effect upon psychologists at the time when it was published. As you read it you should be able to see why.

The third selection, by Lange, uses the twin-method in the study of the tendency to become criminal. In his investigation Lange found strong evidence for believing that to some extent one may be born a criminal. This frightening and still controversial idea is obviously antithetical to Watson's approach. It provides a striking and important contrast.

The fourth selection, again taken from Watson's book, looks at the question of whether we should think of humans as importantly con-

trolled by "instinct." As you might guess, Watson rejects the notion vigorously.

The fifth selection is taken from a book by A. Scheinfeld about identical twins and, as did Lange before him, Scheinfeld finds important evidence for the influence of one's genetic makeup.

Sackett's article, the next in part one, looks at the same basic question in a new way. Sackett reasons as follows: If most knowledge is acquired as a result of our experience in the environment, then a monkey who is raised without ever having the opportunity to see another monkey or even a mirror image of himself should have no real idea of what a monkey's face looks like. On the other hand, if this important knowledge of how to recognize the face of another monkey were somehow genetically coded, then the monkey might "know what another monkey looks like" even though he had never seen one. He might have been *born* knowing how to recognize fellow members of his species. Sackett's experiment gives a clear answer to this question.

In the final article Neal E. Miller looks at a wide variety of evidence from his own laboratories and others in summing up the whole range of genetically determined, and culturally determined aspects of behavior. Miller speaks with considerable authority and his overall position is one commonly held today. Both kinds of influences undoubtedly are vitally important. The task for science at this point is to find out which kinds are important where, and whether the limits which seem to be imposed are absolute or merely relative. The answer is not simple, but it does suggest that both sides were at least partially right all along.

INTRODUCTION

The nature-nurture controversy cuts across a great many questions in psychology. One of the most fascinating areas in which this controversy has been active is in the area of human sexuality. Ever since the time of Freud, it has been widely believed that humans must contend with "a fairly intense, high pressure drive . . . that constrains the individual to seek sexual gratification. . . ." It has been widely accepted that all humans must contend with a powerful innate libidinous "energy" which will influence their behavior from cradle to grave. The following article by Simon and Gagnon reopens the whole question. The authors contend that sexual behavior is greatly and perhaps even primarily determined by socio-cultural influences, and that the innate "programming" may be less specific than has often been assumed. As you read through the various arguments presented in this article, try to relate them to the following question: To what extent is human sexuality influenced by innate biological constraints, and to what extent is it the result of social learning?

ON PSYCHOSEXUAL DEVELOPMENT

William Simon
John Gagnon

Erik Erikson (1963) has observed that, prior to Freud, "sexology" (a wretched and discredited term) tended to see sexuality as suddenly appearing with

Reprinted with permission, from Handbook of Socialization Theory and Research, edited by D. A. Goslin, Rand McNally & Company, 1969, pp. 733-52.

the onset of adolescence. From Erikson's point of view, Freud's discovery of infantile and childhood expressions of sexuality was a crucial part of his contribution. Libido—the generation of psychosexual energies—was now viewed as a fundamental element of the human experience from its very

inception, at the latest beginning with birth and possibly prior to birth. Libido was conceived as something essential to the organism, representing a kind of biological constant with which forms of social life at all levels of sociocultural organization and development, as well as personality structure at each point in the life cycle, had to cope.

In Freud's view, the human infant and child behaved in ways that were intrinsically sexual and that remained in effective and influential continuity with later forms of psychosexual development (Freud, 1953; Chodoff, 1966). Implicit in this view was the assumption that the relation between available sexual energies and emergent motives and attachments would be complex—but direct. In some aspects of psychoanalytic thinking, both adolescent and adult sexuality were viewed as being in some measure reenactments of sexual commitments developed, learned, or acquired during infancy and childhood.

This view presents both an epistemological and sociolinguistic problem. Freud's language was the language of adult sexual experience imposed upon the "apparent" behavior and "assumed" responses, feelings, and cognitions of infants and children. Acts and feelings are defined as sexual, not because of the actor's sense of the experience, but because of the meanings attached to those acts by adult observers or interpreters whose only available language is that of adult sexual experience (Schachtel, 1959). However, the assumptions implicit in the adult terminology with which such behaviors are perceived and considered must be approached with caution. The dilemma arises from the problem of distinguishing between the sources of labeling specific actions,

gestures, or bodily movements as sexual. For the infant playing with his penis, the activity is not sexual in the sense of adult masturbation, but merely diffusely pleasurable in the same manner as are many other activities. The external observer imputes to the child the complex set of states that are generally associated with physically homologous adult activities. It is only through the processes of maturing and learning these adult labels that the child comes to masturbate in the fullest sense of the word. It is in the process of converting external labels into internal capacities for naming that the activities become more precisely defined and linked to a structure of sociocultural expectations and needs that define what is sexual. To suggest that infant or childhood genital play is prototypical of or determines adult patterns is to credit the biological organism with more "natural" wisdom than we normally do in areas where the biological and sociocultural intersect. Undeniably, sexuality is rooted in biological processes, capacities, and possibly even needs. But admitting this in no way provides for a greater degree of biological determinism than is true of other areas of corresponding intersection. Indeed, the reverse may be true: the sexual area may be precisely that realm wherein the superordinate position of the sociocultural over the biological level is most complete.

It is important to note here that the very diffuse quality of most of pre-adolescent experience poses a number of major problems, one of which is the extreme difficulty in getting accurate data. Part of the problem centers on a kind of faulty recall which is not rooted merely in inaccurate memories. This is a source of error pointed out in the existentialist insight that rather

than the past determining the character of the present, it is possible that the present significantly reshapes the past, as we reconstruct our autobiographies in an effort to bring them into greater congruence with our present identities, roles, situations, and available vocabularies. The other part of the problem results from attempting to gather data from children who when interviewed are ill-equipped to report upon their own internal states or from adults who are asked to report on periods of life when complex vocabularies did not exist for them. The problem is attempting to determine what is being felt or thought when confronted with organisms whose restricted language skills may preclude certain feelings or thoughts.

Essential to the perspective of this chapter is the assumption that with the beginnings of adolescence—and with the increasing acknowledgement by the surrounding social world of an individual's sexual capacity—many novel factors come into play, and that an over-emphasis upon a search for continuity with infant and childhood experiences may be dangerously misleading. In particular, it may be a costly mistake to be overimpressed with preadolescent behaviors that appear to be manifestly sexual. In general, it is possible that much of the power of sexuality may be a function of the fact it has frequently been defined as powerful or dangerous. But this overenriched conception of sexual behavior (to the degree that it occurs) must largely follow upon considerable training in an adult language which includes an overdetermined conception of sexuality. Thus it does not necessarily follow that the untrained infant or child will respond as powerfully or as complexly to his

own behaviors that appear to be sexual to an adult observer as will that observer.

Of all the experiences of the pre-adolescent, it would appear that the most influential in the determination of later psychosexual development is that large and complex cluster of elements that come under the heading of sex-role learning. It is perhaps better to label this process "gender-role" learning, underscoring thereby its indirect link to sexuality in many, if not most of its aspects. In other words, we may find that the development of an emerging sense of self is significant in the determination of later sexual commitment or capacities, but that this sense of self *does not* itself derive, in most cases from pre-adolescent sexual or even near-sexual experiences.

For the purposes of this chapter, we reject the unproven assumption of the "power" of the psychosexual drive as a fixed biological attribute. We feel there is little evidence to suggest that such a "drive" need find expression in specific sexual acts or categories of sexual acts. More importantly, we reject the even more dubious assumption that sexual capacities or experiences tend to translate immediately into a kind of universal knowing or wisdom; the assumption that sexuality possesses a magical ability that allows biological drives to seek direct expression in psychosocial and social areas in ways that we do not expect in other biologically rooted behaviors. This assumption can be seen in the psychoanalytic literature, for example, in which the child who views the "primal scene" is seen on some primitive level as intuiting its sexual character. Also, the term latency, in its usage by psychoanalytic theorists, suggests a period of integration by the child of

prior intrinsically sexual experiences and reactions which reduces adolescence, on this level, to a mere management or organization on a manifest level of the commitments and styles already prefigured, if not preformed, in infancy and childhood experiences.

In contradistinction to this tradition, this chapter adopts the view that the point at which the individual begins to respond in intrinsically sexual ways, particularly in terms of socially available or defined outlets and objects, is a time during which something occurs which is discontinuous with previous "sexual experience" (whatever that might mean). Further, at this point in the developmental process, both seemingly sexual and seemingly nonsexual elements "contest" for influence in complex ways which in no respect assure a priority for those elements that are apparently sexual in character.

The prevailing image of the sexual component in human experience is that of a fairly intense, high-pressure drive (except, perhaps, during later stages of the life cycle) that constrains the individual to seek sexual gratification either directly or indirectly. This is clearly present in the Freudian tradition. A similar position is observable in more sociological writings. This is apparent, for example, in the thinking of Kingsley Davis for whom sex is also a high-intensity, societal constant that must be properly channeled lest it find expression in behaviors which threaten the maintenance of collective life (Davis, 1961; Durkheim, 1951).

Our sense of the available data suggests a somewhat different picture of human sexuality, one of generally lower levels of intensity or, at least, greater variability in intensity. More clearly evident among females than males (see below), there are numerous

social situations or roles where reduction and even elimination of sexual activity is managed by greatly disparate populations with little evidence of direct corollary or compensatory intensification in other spheres of life (Gagnon & Simon, 1968a). It is possible that, given the historical nature of human societies, we are victim to the needs of earlier social orders. For earlier societies it may not have been a matter of having to constrain severely the powerful sexual impulse in order to maintain social stability (particularly in family life), but rather a matter of having to invent an importance for sexuality not only to insure high levels of reproductive activity but also to provide a socially available reward that might be placed at the service of many social ends. A part of the legacy of Freud is that we have all become relatively adept at seeking out the sexual ingredient in many forms of nonsexual behavior and symbolism. What we are suggesting is in essence the now three-decade old insight of Kenneth Burke (1935) to the effect that it is entirely plausible to examine sexual behavior for its capacity to express and serve nonsexual motives.

For us, then, sexual behavior is socially scripted behavior and not the masked or rationalized expression of some primordial drive. The individual learns to be sexual as he or she learns sexual scripts, scripts that invest actors and situations with erotic content. One can easily conceive of numerous social situations in which all or almost all the ingredients of sexuality are present, but which remain nonsexual in that not even sexual arousal occurs (Simon & Gagnon, 1967a). Thus combining such elements as desire, privacy, and a physically attractive alter of the appropriate sex, the probability of something

sexual occurring will, under normal circumstances, remain exceedingly small until either one or both actors organize these elements into an appropriate script. The very concern with foreplay in sexual behavior suggests something of this order. From one point of view, foreplay might be defined as progressive physical excitement or what the authors (Gagnon & Simon, 1968b) have elsewhere referred to as the "rubbing of two sticks in order to get a fire going" model. From another point of view, this activity may be defined as a method of eroticizing the body and the activity, as a method for transforming mute, inarticulate gestures and motions into a sociosexual drama.

Lastly, in these introductory comments we might continue to belabor the issue of sociocultural dominance by making some preparatory distinctions, distinctions that we would have to make in any event. Psychosexual development, while a universal component in the human experience, certainly does not occur with universal modalities. Even ignoring the striking forms of cross-cultural variability (Ford & Beach, 1951), we can observe striking differences within our own population, differences that, given the relatively low level of thinking and the sparsity of research in the area, appear to require not a description of psychosexual development, but descriptions of different developmental processes characterizing different segments of the population. The most evident of these are the large number of important differences between observable male and female patterns of sexual behavior (Maccoby, 1966). This particular difference may in some respects be partly attributable to the role played by the biological substratum (Hamburg & Lunde, 1966; Young, Goy & Phoenix, 1964). We

have to account not only for the gross physiological differences and the different roles in the reproductive process that follow from these physiological differences, but must also consider differences in hormonal functions at particular ages. Yet, while our knowledge of many of the salient physiological and physiochemical processes involved is far from complete, there is still little immediate justification for asserting a direct causal link between these processes and specific differential patterns of sexual development observed in our society. The recent work of Masters and Johnson (1966), for example, clearly points to far greater orgasmic capacities on the part of females than males; however, their concept of orgasm as a physiological process would hardly be a basis for accurately predicting rates of sexual behavior. Similarly, within each sex, important distinctions must be made for various socioeconomic status groups whose patterns of sexual development will vary considerably, more impressively for males than for females (Kinsey, 1948). And with reference to socioeconomic status differences, the link to the biological level appears even more tenuous, unless one is willing to invoke the relatively unfashionable conceptual equipment of Social Darwinism. These differences, then, not only suggest the importance of the sociocultural elements and social structure, but also stand as a warning against too uncritical an acceptance of unqualified generalizations about psychosexual development.

Childhood

Obviously Erikson is correct in agreeing with Freud: we do not become fully sexual all at once. There is significant continuity with the past;

nevertheless, continuity is not causality. Even in infancy experiences can occur that will strongly influence later sexual development. However, such experiences will, in all likelihood, be influential not because of their essentially sexual character, but because of their general influential character; that is, they probably influence many more things than just sexual development. It is also possible to talk about the kinds of experience that give rise to fixation, but such experience is probably as unusual as it is traumatic and can hardly be the basis for general theories. One has difficulty conceiving of situations in infancy—or even early childhood—that can be linked to psychosexual development on a level more specific than that of potentiation.

In infancy and in the interrelationships that are part of infant care, we can locate many of the pre-verbal experiences that are preparation for the development of verbal capacities which, in turn, will bind the child to the social world. In this period we can locate some of the experiences—perhaps only sensations—that will help bring about a sense of the body and its capacities for pleasure and discomfort, and we can also locate the experiences that will influence the child's ability to relate to other bodies. The key term remains potentiation: it is possible that through these primitive experiences, ranges are being established, but these ranges are sufficiently broad and overlapping that little can follow by way of specification except through the dubious route of tortured reconstruction. Moreover, if these are profoundly significant experiences to the child—and they may well be—they stand not as expressions of biological necessity nor the inherent wisdom of the body, but as ex-

pressions of the earliest forms of social learning (Gewirtz, 1961).

Unlike the period of infancy, where at best we are confronted with a universe of unnamed gestures, it is possible to conceive of activity in childhood that appears to be explicitly sexual to the observer. About half of all adults report having engaged in some form of "sex play" as children (Kinsey, 1948, 1953). And conceding a certain role to faulty recall or repression, the proportion having behaved this way may be as much as half again as large. The crucial and almost impossible question remains: What does it mean to the child? It is likely that some part of the adult definition of the activity is realized by the participants, but one suspects—as in much of childhood role-playing—that their sense of the adult meaning is fragmentary and ill-formed. It is clear that some learning about the adult world's judgments of the sexual occurs as is indicated by the fact that a high proportion of adults recall that they were concerned, while engaging in childhood sex play, over the possibility of being "found out." However, it may not be the content of sex play activity—as if they were responding to their nascent sexuality—but the mystery that enchants the child (Freud, 1964). Stated differently, the child may be assimilating external bits of information about sex for which, at the time, there are no real internal correlates.

For some small number of persons sexual activity does occur during preadolescence. Probably the largest group are those who become involved as objects of adult-initiated behavior. But for most of them, little seemingly follows. Lacking appropriate sexual scripts the experience remains unassimilated, except per-

haps for those who derive a "meaning" of the experience from a subsequent clinical situation. For some it is equally clear that a severe reaction does follow from falling "victim" to the sexuality of some adult figure, but here it is debatable whether this reaction follows from the sexual act itself or from the social act, the tone, and the intensity of the reactions by others (Gagnon, 1965). In short, relatively few preadolescents become truly sexual, that is, become sexually active. For preadolescent females more often than males, this sexual activity is not immediately related to internal states—capacities for sexual arousal or sexual gratification—but to an instrumental use of sexuality to achieve nonsexual goals and gratifications. This "seductive" preadolescent female, for all her statistical rarity, may represent a significant adumbration of a more general pattern of psychosexual development: a process wherein a commitment to sociosexuality precedes a commitment to sexuality. Among sexually active preadolescent males, behavior appears to be more intrinsically sexual and is associated with a high frequency of subsequent deviant adaptations, most often homosexual (Gebhard et al., 1965).

Of considerable importance is the internalization of values—or the images that powerfully stand for values—that may not be directly or exclusively referential to sexual matters. These values or value images will constitute aspects in the construction of sexual scripts giving rise to senses of the evil, the extraordinary, and the erotic. Despite our present capacity as a society to generate high levels of public discourse about sexual matters, it is probably not unreasonable to assert that learning about sex in our society is learning about guilt; conversely, learn-

ing how to manage sexuality constitutes learning how to manage guilt. An important source of guilt is the imputation by adults of sexual capacities or qualities in children that the children may not have, but that result in—however imperfectly—children learning to act as if they had such capacities or qualities. For example, at what age do girls learn to sit with their knees together, and when do they learn that the upper part of the torso must be hidden, and what of a sexual nature do they learn from all this? Childhood learning of major themes which establish sex or gender role identities is of critical importance. Much of what appears under the heading of sex-role learning involves elements that are remote to sexual experience or that become involved with sexuality only after the latter has become salient (Sears, 1965). Here the meanings and postures of masculinity and femininity are rehearsed and assimilated in many nonsexual ways. Here, also, the qualities of aggression, deference, and dominance needs, which Maslow (1939, 1942)—however imperfect his data—persuasively argues are strongly implicated in the organization of sexual styles, are initially rehearsed, experimented with, and assimilated.

Kagan and Moss (1962) report similar findings: aggressive behavior is a relatively stable aspect of male development and dependency is a similar characteristic in female development. Significant appearances of aggressive behavior by females tended to occur most often among females from well-educated families, families that tended to be more tolerant of deviation from sex-role standards. They also find, of particular interest, that ". . . interest in masculine activities for age 6 to 10 was a better predictor of adult sexual-

ity than was heterosexual activity between 6 and 14." Curiously, they also report that "it was impossible to predict the character of adult sexuality in women from their preadolescent and early adolescent behavior." This the authors attribute to a random factor, that of an imbalance of sex ratios in the local high school and, more significantly, to the fact that "erotic activity is more anxiety-arousing for females than for males" and that "the traditional ego ideal for women dictates inhibition of sexual impulses."

This concept of the importance of early sex-role learning for male children may be viewed in two ways. From one perspective elements of masculine-role learning may be seen as immediately responsive to—if not expressive of—an internal sexual capacity. From another perspective, we might consider elements of masculine identification merely as a more appropriate context within which the mediation of the sexual impulse—which becomes more salient with puberty—and the socially available sexual scripts can occur. Our bias, of course, is towards the latter.

The failure of sex-role learning to be effectively predictive of adult sexual activities, noted by Kagan and Moss, also may lead to alternative interpretations. Again, from one perspective, where sexuality is viewed as a biological constant for both women and men, one can point to the components of female-role learning that facilitate the successful repression of sexual impulses. The other perspective or interpretation suggests differences in the process not of handling sexuality, but of learning how to be sexual, differences between men and women that will have consequences for both what is done sexually as well as when it is done. Our thinking, once again, tends

towards the latter. This position is supported by some recent work of the present authors (Simon & Gagnon, 1967b, 1967c) on female homosexuals, where it is observed that patterns of sexual career management (e.g., the timing of entry into actual sexual behavior, entry into forms of sociosexual behavior, onset and frequency of masturbation, number of partners, reports of feelings of sexual deprivation, etc.) were for lesbians almost identical with those of heterosexual women. Considering what was assumed to be the greater salience of sexuality for the lesbian—her commitment to sexuality being the basis for her entry into a highly alienative role—this seemed to be a surprising outcome. What was concluded was that the crucial operating factor was something that both heterosexual and homosexual women share: the components of sex-role learning that occur before sexuality itself becomes significant.

Social class differences also appear to be significant, although both in the work of Kinsey (1948, 1953) and that of Kagan and Moss they appear as more important factors for males than for females. Some part of this is due to aspects of sex-role learning which vary by social class. Differences in the legitimacy of expressing aggression or perhaps merely differences in modes of expressing aggression come immediately to mind (Sears, 1965; Biller & Borstelmann, 1967). Another difference is the degree to which sex-role models display a capacity for *heterosociality*. The frequently noted pattern of the sexual segregation of social life among working-class and lower-class populations may make the structuring of later sexual activity, particularly during adolescence, actually less complicated (Rainwater, 1966; Simon & Gagnon, 1966). Another as-

pect of social class differences is the tolerance for deviation from traditional attitudes regarding appropriate sex-role performances. Clearly, tolerance for such deviance is positively associated with social class position, and it may well stand in a highly complex and interactive relationship to capacities for heterosocial activities.

We have touched here upon only a few of the potentially large number of factors that should be related to important social class differences and to the processes of psychosexual development. In general, even during this relatively early period of life, complex elements of the ego begin to take form, including the crude outlines of what might be called a repertoire of gratifications. It seems rather naive to conceive of sexuality as a constant pressure upon this process, a pressure that has a particular necessity all its own. For us, this crucial period of childhood has significance not because of what happens of a sexual nature, but because of the non-sexual developments that will provide the names and the judgments that will condition subsequent encounters with sexuality.

Adolescence

Adolescence is a period with ill-defined beginning and end points. There is variation in its general definition, still greater variability in the application of these definitions, and the situation is additionally complicated by the variability of developmental rates of specific individuals. However ill-defined this period may be, the beginning of adolescence marks the time when society, as such, first acknowledges the sexual capacity of the individual. His or her training in the postures and the rhetoric of

the sexual experience will now begin to accelerate. Most importantly, the adolescent will start to view others in his or her immediate environment—in particular, peers, but also some adults—as sexual actors and will find confirmation of this view in the definitions of others toward these actors.

As has already been indicated, for a number of individuals sexual activity begins prior to adolescence, or some portion of children begin engaging in adolescent sexual behavior before they are defined as adolescents. Thus, Kinsey (1948, 1953) reports that by age 12 about a tenth of his female sample and a fifth of his male sample had already experienced orgasm through masturbation. (There is some evidence that early entry into sexual activity is associated with alienative adjustments in later life, but this may not be a function of sexual experience per se so much as the consequence of having fallen out of the more modal socialization patterns and, as a result, having to run greater risks of not receiving appropriate forms of social support.) But this is still an atypical preface to adolescence. For the vast majority, aside from relatively casual childhood sex play and the behaviors that post-Freudians view as masked sexuality, movement into sexual experience which the actor defines and accepts as such begins with the passage into adolescence. Even for persons with prior sexual experience, the newly-acquired definition of their social status as adolescents qualitatively alters the meaning of both current and prior sexual activity; they must now integrate such meanings in more complex ways, ways that are related to both larger spheres of social life and greater senses of self. For example, it is not uncommon during the transitional period between child-

hood and adolescence for both males and females to report arousal and orgasm while engaging in many kinds of physical activity that are not manifestly sexual—climbing trees, sliding down bannisters, or other forms of activity where there is genital contact—without it being defined as sexual by the adolescent (Kinsey, 1948). Indeed, in many such cases there may not even follow subsequent self-explorations in order to achieve some repetition of what was, in all likelihood, a pleasurable experience.

The onset of adolescent sexual development, which really represents the beginning of what will be an adult commitment to sexuality, will be somewhat disjunctive with past experience in most cases. As we have suggested, not only are future experiences to occur in much more complex situations, but also for the first time the more explicit social implications of sexual activity will further complicate matters (Reiss, 1960). The need to manage sexuality, following from a growing sense of having a sexual status, will derive not only from the intrinsic attractions of the sexual experiences for some, but from the increasingly important role sexuality will play in the conduct of both heterosocial and homosocial relationships.

The onset of adolescence demands a separate consideration on our part of the developmental process for males and females. The one thing both genders will share at this point is the reinforcement of their competence in their new sexual status by the occurrence of a dramatic biological event: for males the discovery of the ability to ejaculate, for females the onset of menstruation. This difference can be made evident in the simplest way by pointing to the separate modal routes

into sexual experience. For males the organizing event that initiates a sexual commitment is the biological event of puberty; within two years of puberty all but a relatively few males have their commitment to sexuality reinforced by the experience of orgasm, almost universally brought about by masturbation (Kinsey, 1948). The corresponding organizing event for females is not a biological but a social event: the arrival at an age that suggests a certain proximity to marriage. In contrast to male masturbation rates, we find that only two-thirds of a female population will report ever having masturbated (and then, characteristically, with much lower frequency), and cumulatively we have to go into the latter part of the third decade of life (the twenties) before a proportion is reached which is comparable to that reached by males at age 16 (Kinsey, 1948, 1953). Indeed, it is significant that about half of the females who masturbate will do so only after having their sexuality reinforced by the experience of orgasm in some sociosexual situation. This contrast between males and females, which we will elaborate below, once again points to a distinction between the developmental process for males and that for females: the movement from sexuality to sociosexuality characteristic of males is reversed for females.

When we turn to a more detailed view of the adolescent male experience, despite the fact that we have up to now worked very hard to establish the dominance of the sociocultural or social-psychological over the biological sphere, we must now briefly reverse course. There is some evidence that the early impulse to sexual expression—again initially through masturbation—is linked to high hormonal

inputs during the period of puberty, producing an organism that, to state it in its simplest terms, is easily "turned on" (Kinsey, 1948). It is not uncommon for young male adolescents to report frequent erections during this period, often without the provocation of erotic preoccupations. This obviously focuses considerable attention on the genitalia. Yet, however powerful the biological origins of the propensity for the young adolescent male to masturbate, the meaning and the organization of the activity tend to occur in terms of social and psychological factors.

On one level masturbation is a guilt and anxiety provoking activity for most adolescent males. And despite changes in the rhetoric of public and semipublic discourse on the topic—a shift from a rhetoric of mental and physical destructiveness to one of vaguely inappropriate or nonsocial behavior—this is not likely to change in the foreseeable future. It may be this very element, however, that permits the sexual experience to generate an intensity of affect that may often be attributed to the powers of sexual capacity itself. This guilt and anxiety does not follow simply from the general societal disapproval of this activity. Rather it appears to derive from several sources, one of which is the difficulty encountered in presenting a sexual self to members of the immediate family, particularly parents, where the sentimental and the erotic still remain ill-defined. Another source of guilt and anxiety stems from the fact that arousal and excitement in sexual activity increasingly derives from vicarious organization of sociosexual activity—fantasies involving doing sexual "things" to others or having others do sexual "things" to oneself—or from the assimilation and

rehearsal of available sexual scripts, many of which involve engaging in proscribed activities or relationships. An additional source of guilt or anxiety centers on the general sanction against masturbatory behavior, an activity which few males are willing to admit to after the period of early adolescence.

Despite guilt and anxiety, which may consequently become part of the more general commitment to sexuality for many men, masturbation remains an extremely positive and gratifying experience, as it constitutes for most males through the middle of adolescence the major source of sexual outlet and is engaged in fairly frequently during this period. This form of introduction to sexuality tends to give rise to a capacity for detached sexual activity, activity where the only sustaining motive is sexual; this may actually be the hallmark of male sexuality in our society.

Of the three sources of guilt and anxiety alluded to above, the first, the problem of managing both sexuality and an attachment to family members, is probably common across social class lines. The other two sources should display marked social class differences. The second, the problem of managing a fairly elaborate and exotic fantasy life, should most typically be a problem for higher social class levels. On lower class levels there is a less frequent use of fantasy during masturbation, and this in turn may be linked to a generalized lack of training and competence in the manipulation of symbolic materials (Kinsey, 1948). Higher social class male adolescents are presumably trained for such competence. Masturbation accompanied by fantasy should in turn reinforce the manipulation of symbolic materials. If we de-empha-

size the physical component in the masturbatory act, it can be conceptualized as an activity in which the actor is, in effect, training himself to invest symbols with affect and to derive gratification from the manipulation of symbols. Successfully doing this results in what is classically the best reward or reinforcer in learning theory terms: it is the immediate, specific, and pleasurable experience of orgasm. It may well be that this behavior—which is something females tend not to engage in—plays a role in the processes by which middle-class males catch up with females in measures of achievement and creativity and then, by the end of adolescence, move out in front. This is, of course, merely a wild hypothesis—the Dr. Krankeit (sic) hypothesis (Kenton, 1958). This primary reliance upon masturbation with fantasy should also have a number of consequences beyond the capacity for a relatively detached sexual performance. One such consequence is a tendency to eroticize large parts of the world, as well as an ability to respond to a wide array of visual and auditory *stimuli*. Moreover, to the extent that Wilhelm Reich (1942) is correct in his assertion that the scripting for masturbatory fantasies has a non- and possibly anti-coital character, we might also expect both a capacity and a need for fairly elaborate forms of sexual activity. Further, in so far as the masturbatory fantasy focuses upon relationships and activities essentially preparatory to the coital act, the masturbatory experience should also reinforce an already developing capacity for sustaining heterosocial activity.

The third source of guilt and anxiety, essentially the unmanliness of masturbation, should be more typically a concern of lower social class male adolescents. Among these lower-

class male adolescents the general pattern of sexual segregation of social life and the relatively narrower range of rewarding social experiences available to them should combine to constrain the adolescent to move into heterosexual relationships sooner than does his middle-class counterpart. The first condition, the sexual segregation of social life, should make it easier for him to gravitate toward a world of casual, if not exploitative, sexual relations; it is easier for him than for the middle-class adolescent to learn that he does not have to love everything he desires. The second condition—the more limited available social rewards and particularly rewards deriving from activities that will be validated by his peers—should lead to an exaggerated concern for masculinity-enhancing behavior leading, in turn, to displays of physical prowess, successful staging of aggressive behaviors, and visible sexual success. The three—physical prowess, aggressive behavior, and sexual success—are, of course, not unrelated and frequently are mutually reinforcing.

Available data suggest, in line with the above, that one of the differences between lower-class and middle-class male adolescents is the phasing of entry into heterosexual forms of sexuality or the phasing of movement from masturbation to coital behavior (Kinsey, 1948). In a sense, the lower-class male is the first to reach “sexual maturity” as defined by those working in an essentially Freudian tradition. That is, these lower-class males are generally the first to become exclusively heterosexual and exclusively genital in orientation. One consequence, however, is that while their sexual activity is almost exclusively heterosexual, it also tends to be more homosocial in character. The audi-

ence to which the lower-class male's sexual activity is directed will tend not to be his female partner, but rather it will be more referential to his male peers (Simon & Gagnon, 1966). Middle-class adolescent males, who will initiate coital activity at a significantly later time, will not be exempt from a need and tendency for homosocial components in their sexual lives, but the complexity of their fantasies as well as their social training in an environment less sexually segregated—a world in which distinctions between masculine and feminine roles are not as clearly drawn—will facilitate an easier withdrawal from homosocial commitments. This difference between the social classes will obviously have important consequences for the management of stable adult relations, a point which we will consider below.

One characteristic that will tend to be common to the male experience during adolescence is that while this stage provides extensive opportunity for developing a sexual commitment in one form or other, there is little training in the management of affectively-charged relations with females. The imagery and rhetoric of romantic involvement is abundantly present in the society, and it is not unlikely that a great deal of this is absorbed by the male adolescent, but it is not likely to be significantly tied to a sexual commitment. To the degree that it is related to sexuality, the connection may be quite inhibiting, as is suggested by the still operative "bad girl-good girl" distinction. This is important to keep in mind as we turn now to a consideration of the female side of this story.

The pattern of female development during adolescence is so resistant to social class variation that one is tempted

to seek explanation in something as relatively immutable as the biological level (Kinsey, 1953). However, we will resist this temptation. The female, to be sure, is not provided with the same biologically-rooted incentive to begin an active sexual commitment; but, by the same token, little evidence exists that there is either a social or a biological inhibitor to the development of such a commitment. In simplest terms, the physical equipment for the generation of sexual pleasure is clearly present by puberty and before (Kinsey, 1953), but it tends not to be used by many adolescent females on any social class level. Masturbation rates are fairly low (and those who do masturbate do so fairly infrequently), arousal from explicitly sexual materials and situations is an infrequent occurrence, and there are exceedingly few reports of a self-conscious feeling of sexual deprivation during this period. Explanations for this low level of female sexual activity or interest are numerous on both functional and historical grounds. The basic element in all of these is the idea that females in our society are not encouraged to be sexual, and, indeed, it is possible that they are strongly discouraged from being sexual. As Rheingold (1964) describes this: where men have only to fear sexual failure, women have to fear failure and success simultaneously. Or, as several people have observed, while the category "bad boy" has many descriptive subcategories, the category "bad girl" tends almost exclusively to describe sexual delinquencies. Clearly, it is both difficult and dangerous for a female to become too committed or too sexually active during adolescence.

Whether this extended period of relative sexual inactivity represents the outcome of successful repression

of an elementary sexual urge or merely represents a failure to have an opportunity to learn how to be sexual, is an important question for consideration. The alternative answers have different implications for how we view the later development of a sexual commitment during late adolescence or postadolescence. The "repression answer" suggests that we approach later activity in terms of processes of de-inhibition by which the female learns to find, in varying degrees, modes of more direct expression of internally experienced feelings. It also requires a quest for the sexually determined aspects of nonsexual behavior. The "learning answer" suggests that women create or invent a capacity for sexual behavior, learning how to be aroused and learning how to be responsive. The latter approach also implies greater flexibility in overall adjustments; unlike the repression view, it makes sexuality something other than a constant that is likely to "break loose" at any point in strange or costly ways. In addition, the learning approach lessens the power of the sexual component as a variable; all at once, there is no necessarily healthy or pathological component to a particular style of sexual activity. Lastly, the appeal for us of this approach is somewhat subjective: it tends to seem less like a projection of male sexuality.

The absence of intrinsically sexual activity by adolescent females does not mean that sexual learning and training fail to occur. Curiously, women who are, as a group, far less sexually active than men, receive far more training in self-consciously conceiving of themselves as being sexual on the object level. This is particularly true for recent age cohorts. On the level of the cosmetic self, fe-

males begin relatively early in adolescence to define attractiveness in at least partially sexual terms. One suspects that the same instrumental approach that marked our preadolescent "seductress" now characterizes larger proportions of adolescent females. Parsons' (1954) language of the wife "using" sex to bind the husband to the familial unit, for all its harsh sound, may be quite accurate. In more general terms, the development of a sexual role appears to involve a need to assign to that role services other than pleasure.

To complete this picture of a lack of symmetry between the sexes, the female appears to be trained in precisely that area for which males during adolescence are least trained and for which they are least expected to display a capacity: intense, affect-laden relationships and the rhetoric of romantic love. When sexual arousal is reported by females during this period, it is more often reported as a response to representations of romantic love than as a response to erotic representations.

The movement into later adolescence and the concomitant increase in opportunities for sociosexual activity can be described as a situation in which males—committed to sexuality and relatively untrained in the rhetoric of romantic love—interact with females who are committed to romantic love and relatively untrained in sexuality. Dating and courtship may well be considered processes in which persons train members of the opposite sex in the meaning and content of their respective commitments. And while data in this area are deficient in many regards, the data that are available suggest the exchange system does not often work smoothly. Thus, as is partly suggested by Ehrmann's (1959)

work and partly by our own present studies of college students, it is not uncommon to find, ironically, that the male suitor frequently becomes emotionally involved with his partner and correspondingly less interested in engaging in sexual activity with her, and that the female, whose appreciation of the genuineness of her suitor's affection allows her to feel that sexual activity is now both legitimate and desirable, becomes more interested in engaging in sexual activity with him. Data from the authors' recent study of college students, now under analysis, demonstrate that this difference in commitment is exemplified in several ways. One such item, dealing with the number of times individual respondents had intercourse with their first partner, shows the mode for males around one to three times, while the mode for females is at ten times or more. Clearly, for females, initial intercourse becomes possible only in relatively stable relationships or relationships involving rather strong bonds.

On a theoretical level, we find the male experience conforming to the general Freudian expectation with males moving from a predominantly sexual commitment to an ability to form cathectic attachments in loving relationships. In effect, this movement is reversed for females with cathectic attachments being, in many cases, a necessary precondition for coital activity. It is not surprising, perhaps, that Freud had great difficulty approaching female sexuality. This "error" in conceptualizing female sexuality—of seeing it either as being similar to male sexuality or as a mirrored image—may derive from the fact that so much of the theory construction in this area was done by males. In Freud's case, we also have

to consider the very conception of sexuality that was essential to most of Victorian Europe—it was an elemental beast that had to be curbed.

In addition, it is particularly important to consider social class differences among females, if only as a way of assessing the possibility of a biological factor producing the above-discussed outcome. As we have already indicated, there are very few such differences, far fewer than may be observed among males. One, however, is particularly relevant to this question—the age of first intercourse. This varies inversely with social class and is strongly associated with similar class differences in age of first marriage (Kinsey, 1953). There is no evident basis for associating such social class differences with biological differences. A second difference, perhaps linked only indirectly to social class, is educational achievement. For this variable a single cutting point appears to separate two distinct populations: women who have attended graduate or professional schools and all other women. The former tend to be the most "successful" sexually—at least if one is willing to accept as a measure of success the relatively crude indicator of the proportion of sexual acts that culminate in orgasm. One possible interpretation of this finding derives from Maslow's (1939, 1942) work: women who survive the academic process and go on for additional training are more likely to be more aggressive and/or have strong dominance needs, and both characteristics are associated with heightened sexual commitments. Another somewhat more general interpretation would argue that in a society which still strongly encourages women to form primary allegiances to roles as wives and mothers, the decision to go

on to graduate school represents something of a deviant adaptation. This adaptation represents, in turn, a failure of, or alienation from, modal female socialization processes. And, in effect, it is faulted socialization which produces both the academic or professional commitment and the sexual commitment.

For both males and females, progressively greater involvement with sociosexuality may be one of the factors that marks the end of adolescence. This is a transition about which little is really known, particularly with reference to noncollege populations. Work currently under way by the authors supported by the National Institute of Child Health and Human Development, is attempting to deal with this problem. Our present feeling about the place of sexuality in the management of this transition is that it plays a significant role. First, on a somewhat superficial level, progressive involvement in sociosexuality is important in family formation or in the entry into roles and role obligations that are more explicitly adult in character. But, perhaps on a more fundamental level, it is possible that sociosexual activity is the one aspect of identity experimentation that we associate particularly with later adolescence, a period in which the psychosocial moratorium that Erikson describes as protecting the adolescent during this period of crises and experimentation fails to operate (Erikson, 1963; Reiss, 1960). This may be partly due to the fact that the society has some difficulty protecting the adolescent from the consequences of that part of his behavior it is not prepared to admit he is engaging in. More importantly, it may be due to the fact that we have, at all age levels, great problems in talking about sexual

feelings and experiences in personal terms which, in turn, make it extremely difficult to get social support for our experiments with our sexual selves (the term "pluralistic ignorance" is perhaps nowhere more applicable than in the sexual area). It may be that these experiments with sexual capacities and identities rank among the first unprotected tests of competence and the quest for a basis for self-acceptance. We suspect that success or failure in the management of sexual identity may have consequences in many more areas of personality development than merely the sexual sphere.

Adulthood

All but very few persons in our society ultimately marry. The management of sexual commitments within a marital relationship characterizes the larger part of postadolescent experience in our society. Once again, it is important to underscore the real poverty of data on this topic. Sexual adjustment presumably plays an important role in overall marital adjustment. This judgment largely derives from studies of broken marriages or marriages that are in trouble, and we really have very little sense of the degree to which sexual problems in troubled or dissolved marriages exceed those found in marriages which remain intact. It is possible that we have assumed an important role for sexuality and the management of sexuality in the maintenance of marital bonds because we have assumed sex itself to be an important part of most people's lives. This may not be true. Particularly after the formation of the marital unit, it is quite possible that sex declines in salience. It may stand as less important than alternative

modes of gratification, or the weight of alternative gratifications may minimize the effects of sexual dissatisfaction. It is also possible that individuals learn to derive sexual gratification from non- or only partially sexual activities. This is not to suggest support for the concept of sublimation, but rather to point out that in the processes that follow marriage, newly learned alternative patterns of gratification may substitute for the sexual.

The main determinant of adult rates of sexual activity in our society is the level of male commitment. While interest in intercourse is highest for males during the early years of marriage, a corresponding peak in coital interest occurs much later in marriage for females (who require longer periods of time to either become de-inhibited or to learn to be sexual—depending upon your point of view). Nonetheless, coital rates in marriage decline steadily through marriage (Kinsey, 1953). This decline, it should be noted, can only be partly attributed to declines in biological capacity on the part of the male. The decrease may derive from many things. In many cases the problem is one of relating sexually to a person whose roles have become complicated by the addition of maternal functions (Freud, 1949). For lower-class males, there is a problem of not receiving homosocial support for marital intercourse, to which we might also add the disadvantage of being less trained in the use of auxiliary materials to heighten sexual interest (Simon & Gagnon, 1966). For the middle-class males, the decline is less steep, owing perhaps to their ability to find sexual stimulation from auxiliary sources—literature, movies, etc. Also operative is a greater capacity for invoking and responding to fantasy. It should be

noted that for about 30 per cent of college-educated males, masturbation continues as a regular source of sexual outlet in marriage and during periods when a wife is available (Kinsey, 1948). To this we might add an additional but unknown proportion who do not physically engage in masturbation but for whom the source of sexual excitement is not just coital activity alone but also the fantasy elements which accompany coital activity. But even for the middle-class male, sexual activity declines in degrees that cannot exclusively be accounted for by changes in the organism. Perhaps it is simply that the conditions under which we learn to be sexual in our society make it extremely difficult to maintain high levels of sexual performance with a single partner over long periods of time. This may remain relatively unimportant in the maintenance of the family unit or even the individual's sense of his own well-being because of the relative unimportance of sexual dissatisfaction or the relatively greater significance of other areas of life.

About half of all ever-married males and a quarter of all ever-married females will engage in extramarital sexual activity at one time or another. For females there is some suggestion of a secular trend toward increases in extramarital activity from the beginning of this century to the early fifties. This is linked to a corresponding generational rise in rates of orgasm during this same period (Kinsey, 1948, 1953). It is possible that the very nature of female sexuality may be undergoing change. Our current data will hopefully shed some light on this. For males there are strong social class differences, with lower-class males accounting for most of the extramarital activity, particularly during the early



years of marriage. This may be a direct reflection of their earlier mode of assimilation of the sexual commitment. As we previously observed, it is difficult for lower-class males to receive homosocial validation from marital sexual activity (unless, of course, it culminates in conception); this is not the case for extramarital activity for which there is abundant homosocial validation.

In general, it is our feeling that sexuality and sexual activity are by and large derivative functions even during adulthood. There are only a few periods in the life cycle at which there are high rates of sexual activity and/or sexual activity that is complicated by passion and high intensity of affective investment. These are usually adolescence in the male, the early and romantic years of marriage for both men and women, and the highly charged extramarital experiences that can be called affairs. Most of the time sex is really a relatively docile beast, and it is only the rare individual who through the processes of self-invention or alienation from the normal course of socialization is prepared to risk occupation, present comfort, wife and children, or the future for the chancy joys of sexual pleasure.

From this point of view it might be more proper to suggest that, contrary to the Freudian point of view that sex manifests itself in some form in other types of conduct or that other conduct is symbolic of sexual conflict, reports of sexual conflict may in fact stand for difficulties in the more conventional zones of life. Thus, the married couple who come for counseling because of a sexual problem may be merely reporting the conventional rhetoric of the society about what they think the sources of a marital difficulty ought to be. Ovesey (1950) reports that homosexual dreams (overtly homosexual,

not merely symbolic) of heterosexual men really relate to occupational problems and that the submissiveness required occupationally appears in the convenient symbolism of the "purported" femininity of homosexual relations. Indeed, many forms of both heterosexual and homosexual acting-out seem to be related to stress reactions to other life situations rather than having specifically sexual motivations. Thus, studies of sex offenders often are overly concerned with the sexual life of the offender when this may merely be the symptom of disorders of other kinds.

Conclusion

It is only a fairly recent development in the history of man that he could begin to conceive of the possibility of social change, that he could begin to understand that his time and place did not represent the embodiment of some eternal principle or necessity, but was only a point in an on-going, dynamic process. For many it is still more difficult to conceive of the possibility of the nature of man himself changing, and particularly changing in significant ways (Van den Berg, 1964). Much of this conservative view of man still permeates contemporary behavioral science. Thus, for many social theories, a view of man as a static bundle of universal needs supplies the necessary stability not available elsewhere in the flux of social life. A conception of man as having relatively constant sexual needs is a necessary part of this point of view. As a contrast to this conservative view, we have attempted to offer a description of sexual development as a variable socio-cultural invention, an invention that in itself explains little and requires much explanation.

References

- BILLER, H. B., & BORSTELMANN, L. J., Maculine development: An integrative review. *Merrill Palmer Quarterly*, 1967, 13, 253-294.
- BURKE, K., *Permanence and change*. New York: New Republic, Inc., 1935.
- CHODOFF, P., Critique of Freud's theory of infantile sexuality. *American Journal of Psychiatry*, 1966, 123, 507-518.
- DAVIS, K., Sexual behavior. In R. K. Merton & R. Nesbitt, *Contemporary social problems*. (2nd Ed.) New York: Harcourt-Brace, 1961.
- DURKHEIM, E., *Suicide*. Glencoe: The Free Press, 1951.
- EHRMANN, W., *Premarital dating behavior*. New York: Holt, 1959.
- ERIKSON, E. H., *Childhood and society*. (2nd Ed.) New York: Norton, 1963.
- FORD, C. F., & BEACH, F. A., *Patterns of sexual behavior*. New York: Harper, 1951.
- FREUD, S., The most prevalent form of degradation in erotic life. *Collected papers*, Vol. 4. London: Hogarth, 1949.
- FREUD, S., Three essays on sexuality. *Complete psychological works*. (Std. Ed.) Vol. VII. London: Hogarth, 1953. Pp. 135-245.
- FREUD, S., Analysis terminable and interminable. *Complete psychological works*. (Std. Ed.) Vol. XXIII. London: Hogarth, 1964. Pp. 216-253.
- GAGNON, J. H., Female child victims of sex offenses. *Social Problems*, 1965, 13, 176-192.
- GAGNON, J. H., & SIMON, W., The social meaning of prison homosexuality. *Federal Probation*, 1968. (a)
- GAGNON, J. H., & SIMON, W., Sex education and human development. In P. J. Fink (Ed.), *Human sexual function and dysfunction*. Philadelphia: F. A. Davis, 1968. (b)
- GEBHARD, P. H., ET AL., *Sex offenders*. New York: Harper & Row, 1965.
- GEWIRTZ, J. L., A learning analysis of the effects of normal stimulation, privation, and deprivation in the acquisition of social motivation and attachment. In B. M. Foss (Ed.), *Determinants of infant behavior*. New York: Wiley, 1961. Pp. 213-290.
- HAMBURG, D. A., & LUNDE, D. T., Sex hormones in the development of sex differences in human behavior. In E. Maccoby (Ed.), *The development of sex differences*. Stanford: Stanford University Press, 1966. Pp. 1-24.
- KAGAN, J., & MOSS, H. A., *Birth to maturity*. New York: Wiley, 1962.
- KENTON, M. (pseud. of Southern, T. & Hoffenburg, M.), *Candy*. Paris: The Olympia Press, 1958.
- KINSEY, A. C., ET AL., *Sexual behavior in the human male*. Philadelphia: Saunders, 1948.
- KINSEY, A. C., ET AL., *Sexual behavior in the human female*. Philadelphia: Saunders, 1953.
- MACCOBY, E. (Ed.), *The development of sex differences*. Stanford: Stanford University Press, 1966.
- MASLOW, A. H., Dominance, personality, and social behavior in women. *Journal of Social Psychology*, 1939, 10, 3-39.
- MASLOW, A. H., Self esteem (dominance feeling) and sexuality in women. *Journal of Social Psychology*, 1942, 16, 259-294.
- MASTERS, W. H., & JOHNSON, V. E., *Human sexual response*. Boston: Little, Brown and Co., 1966.
- OVESEY, L., The homosexual conflict: An adaptational analysis. *Psychiatry*, 1950, 17, 243-250.
- PARSONS, T., The kinship system of the

- contemporary United States. In *Essays in sociological theory*. (Rev. Ed.) New York: The Free Press, 1954.
- RAINWATER, L., The crucible of identity: The Negro lower class family. *Daedalus*, 1966, 95, 172-216.
- REICH, W., The function of the orgasm. In *The discovery of the Orgone*. Vol. I. New York: Orgone Institute Press, 1942.
- REISS, A. J., JR., Sex offenses: The marginal status of the adolescent. *Law and Contemporary Problems*, 1960, 25, No. 2.
- RHEINGOLD, J. C., *The fear of being a woman*. New York: Grune & Stratton, 1964.
- SCHACHTEL, E., *Metamorphosis*. New York: Basic Books, 1959.
- SEARS, R. R., Development of gender role. In F. A. Beach (Ed.), *Sex and behavior*. New York: Wiley, 1965. Pp. 133-163.
- SIMON, W., & GAGNON, J. H., Heterosexuality and homosociality: A dilemma of the lower class family. 1966. (Mimeographed)
- SIMON, W., & GAGNON, J. H., Pornography, raging menace or paper tiger? *Trans-Action*, 1967, 4, 8, 41-48. (a)
- SIMON, W., & GAGNON, J. H., The lesbians: A preliminary overview. In J. H. Gagnon & W. Simon (Eds.), *Sexual deviance*. New York: Harper & Row, 1967. (b)
- SIMON, W., & GAGNON, J. H., Femininity in the lesbian community. *Social Problems*, 1967, 15, 212-221. (c)
- VAN DEN BERG, J. H., *The changing nature of man*. New York: Dell, 1964.
- YOUNG, W. R., GOY, R., & PHOENIX, C., Hormones and sexual behavior. *Science*, 1964, 143, 212-218.

INTRODUCTION

The following selection is taken from Watson's landmark book *Behaviorism*, which had an enormous effect, both positive and negative, upon the development of psychology. Watson's book was first published in 1924, but you will see that this brief selection could have been written yesterday. Experts still disagree violently as to whether psychopathology should be described as mental disease, or as behavior disorder. This question comes up again later, in Part Five, *The Control of Human Behavior*. Watson's ideas are very instructive not only because he writes well but especially because he has deliberately taken a polar position: the position that hereditary constraints are negligible and that all important changes in human development are due to environmental forces. Watson's book had a profound affect upon psychologists at the time when it was published. As you read this brief excerpt, you should be able to see why.

ARE THERE SUCH THINGS AS MENTAL DISEASES?

John B. Watson

I know that all of this more or less vague discussion about the analyst and the physician raises some definite questions which may be phrased as follows: Isn't there such a thing as a mental disease? If so, what is it like and how do you cure it?

As long as there is the misconception going the rounds that there is such a thing as the mental, I suppose there will be mental diseases, mental symptoms and mental cures. I view the whole matter otherwise. I can only roughly outline my own viewpoint. Sickness of personality, or behavior illnesses, behavior disturbances, habit conflicts, are terms which I should like to use in place of mental disturbances, mental diseases and the like. In many of the so-called psychopathological disorders ("functional psychoses," "functional neuroses,") there are no organic disturbances of sufficient gravity to account for per-

sonality disturbance. There may be no infections, no lesions anywhere, no absence of physiological reflexes (as there often is when there are organic diseases). And yet the individual has a sick personality. His behavior may be so badly disturbed or so involved that we call him insane (a purely social classification) and have to commit him temporarily or permanently.

No one has yet given us a sensible classification of the various types of behavior disturbance that are present in a social structure such as ours. We hear of manic depressive insanity, anxiety neuroses, paranoia, schizophrenia and many others. To me as a layman these divisions mean nothing. I know in general what is meant by appendicitis, cancer of the breast, gall stones, typhoid fever, tonsillitis, tuberculosis, paresis, brain tumor and even cardiac insufficiency. I know in general something about the organisms present when any are present, the kind of tissue attacked and the general course of the disease and I can understand the physician when he tells me more about them. Yet when the psychopathologist tries to tell me about a "schiz" or a "homicidal mania" or

Reprinted with permission, from Behaviorism by John B. Watson, W. W. Norton & Company, Inc. Copyright 1924, 1925 by the Peoples Institute Publishing Co., Inc. Copyright renewed 1952 by John B. Watson. Revised edition Copyright 1930 by W. W. Norton & Company, Inc. Copyright renewed 1958 by John B. Watson, pp. 297-300.

an "hysterical" attack, I have the feeling, which has grown stronger with the years, that he doesn't know what he is talking about. And I think the reason he doesn't know what he is talking about is that he has always approached his patients from the point of view of the *mind* rather than from that of the way the whole body behaves and the genetic reasons for that behavior. The past few years have unquestionably seen considerable progress in this direction.

To show the needlessness of introducing the "conception of mind" in so-called mental diseases, I offer you a fanciful picture of a psychopathological dog (I use the dog because I am not a physician and have no right to use a human illustration—I hope the veterinarians will pardon me!). Without taking any one into my counsel suppose I once trained a dog so that he would walk away from nicely-ground, fresh hamburger steak and would eat only decayed fish (true examples of this are now at hand). I trained him (by use of the electric shock) to avoid smelling the female dog in the usual canine way—he would circle around her but would come no closer than ten feet (J. J. B. Morgan has done something very close to this on the rat). Again, by letting him play only with male puppies and dogs and punishing him when he tried to mount a female, I made a homosexual of him (F. A. Moss has done something closely akin to this in rats). Instead of licking my hands and becoming lively and playful when I go to him in the morning, he hides or cowers, whines and shows his teeth. Instead of going after rats and other small animals in the way of hunting, he runs away from them and shows the most pronounced fears. He sleeps in the ash can—he fouls his own bed, he urinates every

half hour and anywhere. Instead of smelling every tree trunk, he growls and fights and paws the earth but will not come within two feet of the tree. He sleeps only two hours per day and sleeps these two hours leaning up against a wall rather than lying down with head and rump touching. He is thin and emaciated because he will eat no fats. He salivates constantly (because I have conditioned him to salivate to hundreds of objects). This interferes with his digestion. Then I take him to the dog psychopathologist. His physiological reflexes are normal. No organic lesions are to be found anywhere. The dog, so the psychopathologist claims, is mentally sick, actually insane; his mental condition has led to the various organic difficulties such as lack of digestion; it has "caused" his poor physical condition. Everything that a dog should do—as compared with what dogs of his type usually do—he does not do. And everything that seems foreign for a dog to do he does. The psychopathologist says I must commit the dog to an institution for the care of insane dogs; that if he is not restrained he will jump from a ten-story building, or walk into a fire without hesitation.

I tell the dog psychopathologist that he doesn't know anything about my dog; that, from a standpoint of the environment in which the dog has been brought up (the way I have trained him) he is the most normal dog in the world; that the reason he calls the dog "insane" or mentally sick is because of his own absurd system of classification.

I then take the psychopathologist into my confidence. He becomes extremely angry. "Since you've brought this on, go cure him." I attempt then to correct my dog's behavior difficulties, at least up to the point where he

can begin to associate with the nice dogs in the neighborhood. If he is very old or if things have gone too far, I just keep him confined; but if he is fairly young and he learns easily, I undertake to retrain him. I use behavioristic methods. I uncondition him and then condition him. Soon I get him to eating fresh meat by getting him hungry, closing up his nose and feeding him in the dark. This gives me a good start. I have something basal to use in my further work. I keep him hungry and feed him only when I open his cage in the morning; the whip is thrown away; soon he jumps for joy when he hears my step. In a few months' time I not only have cleared out the old but also have built in the new. The next time there is a dog show I proudly exhibit him, and his general behavior is such an asset to his sleek, perfect body that he walks off with the blue ribbon.

All this is an exaggeration—almost

sacrilege! Surely there is no connection between this and the poor sick souls we see in the psychopathic wards in every hospital! Yes, I admit the exaggeration, but I am after elements here. I am pleading for simplicity and ruggedness in the building stones of our science of behavior. I am trying to show by this homely illustration *that you can by conditioning not only build up the behavior complications, patterns and conflicts in diseased personalities, but also by the same process lay the foundations for the onset of actual organic changes which result finally in infections and lesions—all without introducing the concepts of the mind-body relation ("influence of mind over the body") or even without leaving the realm of natural science.* In other words, as behaviorists, even in "mental diseases" we deal with the same material and the same laws that the neurologists and physiologists deal with.

INTRODUCTION

Lange's study used the "twin-method" in the study of the tendency to become criminal. This method takes advantage of the fact that occasionally nature produces a pair of men with identical genetic makeup—identical twins, also called monozygotic twins. (Monozygotic means simply that both individuals in a pair of identical twins originated in a single cell, or zygote, formed at conception. This cell split into two identical cells each of which eventually matured into a complete being—one of the identical twins.) This happy circumstance means that we can look analytically at the extent to which the future behavior of these two individuals is similar even when they have grown up under very different conditions. To the extent that such similarity is found, and to the extent that there is more similarity for identical twins than for nonidentical twins, we can conclude that the genetic influences are significant and are strong enough to overcome differences in environment.

As you will see, Lange is confident that he has found evidence that some people are cursed with a predisposition toward criminal behavior. Lange believes that some people, because of their genetic makeup, are more likely to become criminals than are others. This frightening and still controversial idea is obviously antithetical to Watson's approach. It provides a striking and important contrast. Those of you who have followed the recent controversy about the "XXY" chromosome, which some people believe indicates a predisposition to criminal behavior, will appreciate that Lange's book, published in 1931, may have been substantially ahead of its time.

CRIME AS DESTINY

Johannes Lange

"It would be an . . . improvement of our present-day justice if every year a lottery were held among all citizens of this free land of ours in order to decide who should go to prison and who should not. The result of the draw would represent just as impartial a form of justice as the one we now have, and would also give each of us an opportunity for once to be punished and reformed as a miserable sinner. God would keep an eye on the matter and would take care that even those who had not drawn prison as their lot would receive what they deserved." Thus writes Lindsey, a judge of the American children's courts, and according to his alleged, though not altogether likely, experiences he ought to know. God, before whom we are all guilty, is his witness, and for those whose life is poised between mercy and hell, Lindsey's words certainly must have a deep meaning. Lindsey also says nothing which they will straightway deny to the many who have unmitigated insight into the phenomena on the edge of their consciousness and into the final bases of their conduct, and who have nevertheless not been able to free themselves.

In Germany as well as in America many earnest people incline to views

Reprinted with permission from Crime as Destiny by Johannes Lange, George Allen & Unwin, Ltd., London, 1931, pp. 27-45.

which come very close to those of the American judge of youth. According to such views the criminal is set apart from the mass of humanity, not by the stock he springs from, but by accidental and difficult circumstances which control his life-history. The blame is chiefly put on experiences in early childhood and unpropitious influences as regards education. In this way society is most to blame for crime and the criminal. He is at bottom a martyr who only needs our keenest sympathy, as well as help, education, and training, to become as good a member of human society as anyone else.

"All things are good as they came out of the hands of their Creator, but everything degenerates in the hands of man." This famous statement of Rousseau's is accepted by more people to-day than was ever the case since it was written.

In other times people thought otherwise, and even today only very few would agree unconditionally with Lindsey. The view that human beings are intrinsically quite different from one another and that the environment given them by fate has only very small importance for the development of personality is another held quite bluntly and also by quite serious people. Galton once used the example of the cuckoo, who sings the same song in all parts of the world, although his

eggs are laid in all kinds of different nests. The "continuous merciless march of the hidden weaknesses in our nature through sickness to death" and also in the spiritual sphere was what Galton's investigations revealed, and he considers it a fair question to ask whether environmental influences have any part at all in the development of personality apart from the fact that they facilitate knowledge and professional training. "One cannot avoid the conclusion that heredity has an enormously greater weight than environmental influences when the differences in the latter do not go beyond those commonly found between persons of the same class and the same country." Galton does not actually refer to the problem of the criminal, but it is indubitable that if his conclusions were correct they would have great importance in the field of criminology.

Such coarsely differentiated points of view are only possible where emotional reactions definitely influence the judgment. That emotional needs must play an important part in the question of how crimes come to be committed is obvious. Moral, religious, ethical, political considerations colour the question before one has even considered the facts. Taken alone, these could give one completely clear and undeniable conclusions without abolishing the contradictions of the various points of view. Galileo denied his incontestable discovery before the religious tribunal, and yet it was as certain as it had been before that the earth was the centre of the universe. But here, in the case of the problem of the criminal, even the facts are not completely clear. We do know numberless details but nothing conclusive.

Thus we know the quantitative re-

lationship between certain social phenomena and crime; for example, the number of crimes against property rises during bad economic periods. Crimes of violence and sexual crimes have a definite yearly and even perhaps weekly curve, due to external causes, some of which we know. Therefore outward influences are certainly not without importance as causes of crime. On the other hand, we know criminal families who have been followed up for several generations and in whom the inclusion of bad as well as good blood is revealed in an unmistakable manner. It has long ago been shown that hardened criminals are not seldom descended from criminal parents or else have numerous criminal relations. We also know that hereditary taints, such as mental diseases, and particularly alcoholism and psychopathic disturbances, are very common in criminals, and that many of them are themselves mentally sick or otherwise abnormal. We know types of so little intelligence and so lacking in all social feeling that sooner or later they simply must come up against the law. Owing to the importance of this fact the chief interest in the investigation of crime has shifted during recent years more and more from the deed to the doer. In spite of this the controversy as to whether unfavourable tendencies or environment are mostly responsible for crime has not been settled. At any rate, Lombroso's dream of "the born criminal," "the natural delinquent," as a special human type has been dreamed to a finish. Nothing remains of it except the sterile fact that a large number of criminals are in some way or another abnormal, just as are many other people who never come into conflict with the law at all.

We have hardly got a definite step

beyond such general observations. However much we may talk, we still do not know enough about the great majority of criminals to make certain of dealing with them successfully.

This situation is not very satisfactory from the point of view of our civilisation. Decisions of the most far-reaching importance are made quite blindly. We erect new legal codes without having tackled the basic questions at all seriously. We are getting rid of the concept of punishment. We take the most comprehensive precautions to safeguard society, we sterilise thousands of criminals, and on the other hand we claim for a number of others protection on the ground of low powers of self-control which make them a danger to society, while we cannot know at all clearly who should be sterilised or who should be permanently protected. We have no means for the simplest collection and investigation of the real basic material which should precede any such measures.

That is the general position to-day, and yet in certain cases, at any rate, the possibility of taking a right or a false attitude to the problem rests with ourselves. Are we going to take a blind risk that in the future justice shall become less and less secure, that criminal tendencies (if there really are such things) shall be propagated without check, or shall we, on the other hand, destroy an irreplaceable hereditary material? Must we decide the question whether or not to abolish the death penalty entirely from the point of view of political expediency? I should have thought that any society would go to the limit of trouble to get this matter cleared up.

Actually the beginnings of reform can be seen in one or two places. Above all, a few educationists have set

themselves to deal with young law-breakers partly through professional interest, partly through sympathy blended with such high ideals and so much self-sacrifice that when they have done all they can they will still put down failure to their own inadequacy rather than that of their charges. "We should be still more careful to educate from the child's point of view and to discover the roots of his tendencies in order to form his character according to his innate possibilities. We are still fumbling about far too much," was the answer given me by an admirable and admired woman when I ventured to criticise some of her work on the ground that she had made too much of inborn tendencies. If such humanists had their way every difficult adolescent would have more than one permanent guide to look after him, even though it were not at all certain that all the trouble taken would result in success. We can certainly not decide to-day whether it will be possible for any given individual to adapt himself successfully to social conditions.

As a preliminary to all intelligent measures we need wide and deep knowledge of criminals. In this field the State, with its huge organisations and materials, can do decisive work. And this task has been begun in a few places. Above all I must mention the Bavarian Ministry of Justice, which in spite of opposition has installed at Degen's instigation the Institute for Criminal Biology at Straubing Prison. Here are collected as many records as possible of criminals which have been thoroughly gone into by prison doctors in order to lay the basis of true knowledge. The number of investigators is a small one and the investigations have to be made in addition to routine work, which takes up almost all of the doctors' time.

There is a dearth of people and means to make full use of the material collected, and with a curious failure to realise the importance of these investigations the State fails to provide money to improve the situation. Yet so much has been done that there is no place for destructive criticism. Those who criticise most loudly can bring up fewest facts to support them. Their own proposals can only be tested by the use of a mass of material which can be procured solely by the methods they attack. But it is a good sign that in spite of all criticism, many States are following the Bavarian example, and in other directions similar experiments are being made.

When we have really got a mass of material thoroughly investigated from all points of view, then we shall be able to deal with criminals with our eyes open. Until then we must go warily. However, there is one way—the way we have taken here—to clear up some basic questions more definitely than was hitherto possible. This is by the use of the so-called Twin method.

The Twin Method

Twins are brothers and sisters born together and who in the majority of cases have grown up together. In the decisive years of development they have had common experiences and have been subject to common educational influences. Apart from pairs consisting of a twin brother and a twin sister, there are pairs of twins of the same sex, but of quite different conception and formation.

The most famous example of one sort of twins of the same sex are Jacob and Esau, the sons of Isaac and Rebecca. They were not only different in appearance—Esau was red-skinned

and rough, Jacob was fair and smooth—they had different voices and a different odour. In their tastes and their characters they differed equally. Esau was a hunter, rough, straightforward, uncontrolled, violent-tempered, and simple, and married wives of a lower class, sometimes against his father's orders; Jacob was a shepherd, smooth-speaking, artful, lying, out for his inheritance, and sly—at the same time tough, moody, and unjust. In choosing his brides he was cautious and self-sacrificing. The brothers were about as different from one another as sibs¹ could possibly be. Twins as different as these in appearance and character are quite common; they are called dizygotic twins, and there is good reason to suppose that they have been born from two ova which were simultaneously fertilised. According to their innate dispositions they resemble or differ from one another about as much as ordinary sibs.

On the other hand, there are twins of the same sex who look so alike that it is almost impossible to tell them apart and whose personalities are also practically identical. According to general opinion, they have both sprung from one and the same fertilised egg and have therefore exactly the same hereditary disposition. A well-known example in literature are the brothers Weidelich in Gottfried Keller's *Martin Salander*. I suspect that they were drawn from life. They were as like as one egg is to another; in their behaviour they were also exactly similar. Independently both committed deceptions at the same time and in the same way. Outwardly they were just a little different according to circumstances.

The first person to think of using

¹ Sib is a collective name for all children born of the same parents.

monozygotic and dizygotic twins to solve the problem of heredity and environment was Galton, who in 1876 published results which are still important today. He chose two different lines of investigation, taking first very similar, i.e. monozygotic, twins in order to see what possible differences later developed between them, and then very dissimilar, i.e. dizygotic, twins in order to find out whether the same outward influences increased their resemblance. He came to the conclusion already mentioned, that heredity has a far greater influence than environment. Twins with the same heredity seldom become different from one another, and if they do the difference is not due to "free-will." This was never denied by Galton's numerous clerical informants. On the other hand, similar experiences did not increase the resemblance between dizygotic twins.

Increasing experience has shown more and more clearly that twins with markedly similar outward appearance agree with one another in a large number of definite characteristics, such as the colour and type of their hair, eye colour, skin colour, distribution of body hair, etc., whereas with dizygotic twins these characteristics are like those of other sibs, i.e. they are very seldom exactly similar, or, if some may be, the whole lot never are. Further, it was found out that monozygotic twins always had the same hereditary diseases or else both would remain free from family troubles, whereas nonidentical twins were in this respect like their other brothers and sisters, i.e. they might both be ill at the same time, but this did not happen more often than one would expect, and mostly only one of the twins fell ill while the other remained well. For it is inborn tendencies which con-

trol the appearance of hereditary diseases, and inborn tendencies are exactly the same in monozygotic twins, but different in the dizygotic pairs. In the case of the latter, in view of their similar parentage about half the innate tendencies are the same.

Now there are diseases which are not exclusively due to heredity but in whose case outward circumstances play an important rôle. With regard to these, monozygotic and dizygotic twins show characteristic differences. The greater the weight of heredity, the more closely each of a given pair of monozygotic twins will resemble one another; as the importance of heredity is reduced the susceptibility of monozygotic twins will resemble that of two or more children born at the same time from separately fertilised eggs.

The resemblances between monozygotic twins do not stop at physical characteristics or disease; there are remarkable resemblances on the mental plane. It is very rare to find human beings who are really identical from this point of view—for example, their writing is generally different—but it has been found that exactly similar attitudes on the part of monozygotic twins towards essential matters, such as a career or marriage, or their general relationship to their environment, are extremely common. Here again similar heredity plays its part. Hereditary nervous and mental troubles affect both of a pair of monozygotic twins almost without exception, or else both escape them.

All this justifies us in using the results obtained with twins in order to examine the question whether, and if so to what extent, experiences of the most different kinds are determined by our inborn tendencies. Siemens especially has systematically and suc-

cessfully used this method. Complete agreement in the case of closely resembling twins backs up the view that heredity is chiefly, or at any rate preponderantly, responsible for one's experience. According to how much the behaviour of hereditarily identical twins differs, by just so much is the influence of heredity in determining our fate diminished.

Use of the Twin Method in Investigating the Causes of Crime

A. GENERAL

It was fairly obvious that the Twin method could be used to determine the influence of innate tendencies in the causation of crime. If the hereditary make-up had no importance, a comparison between monozygotic and dizygotic pairs of twins ought to show no differences. Agreement between the behaviour of monozygotic twins would be in accordance with the importance of heredity. A lack of agreement in the case of twins with identical heredity would enable us to estimate the importance of environmental incentives to crime. Finally, we could compare the behaviour of dizygotic twins with that of other brothers and sisters. If in comparison with other brothers and sisters twins with a different heredity showed closer agreement as far as crime was concerned, the importance of the environmental influences would be shown to be increased in proportion, as the latter can only be considered exactly similar in the case of those who have grown up together.

B. MATERIAL

Twins are not very common; roughly speaking, one pair is born to

every eighty ordinary births. In addition, about 40 per cent. of twins die, a far higher percentage mortality than that of average children, owing to their weakness in infancy. The only twins considered suitable for our investigation were those who were both old enough to be prosecuted by the law; for it very often happens that one or the other twin dies young. Brother and sister pairs did not count for us, as it is well known that the differences between the sexes, as far as crime is concerned, are very great. Thus we had to find one of a pair who had been imprisoned and whose other twin was still living, was old enough to be able to come in conflict with the law, and was of the same sex.

The material was provided by the records of the Institute for Criminal Biology. Furthermore, at our request the Bavarian Ministry of Justice ordered that all prisoners in Bavarian prisons who were twins should be reported and examined from the point of view of Criminal Biology. In addition we also asked for such prisoners as had twins among their brothers and sisters who might also be of an age to be sentenced. Finally, I looked among the psychopathic patients of the General Department of the German Institute for Psychiatry for twins who had been imprisoned. I also asked all twins whom I met in the course of my hospital duties for criminal records. All those twins who fulfilled the above-mentioned conditions were taken as subjects. First of all, the criminal records of the subjects themselves and of their twin were examined. Then I interviewed those twins who were still in prisons. We then went into the degree of resemblance and the life-stories of all the subjects. Police records and sentences were, of course, examined at the same time.

These were put at my disposal as well as other ordinary official documents. On various other pretexts I got some of the criminal twins to come for interviews to my consulting-room. Others I visited at their homes in different cities. My former colleague, Dr. Grüber, undertook particularly detailed researches at my request. I am also very grateful to Professor Ewald and Dr. Faltlhauser in Erlangen, Frau Dr. Schmidt-Kraepelin in Ludwigshafen, Fräulein Dr. Emy Metzger in Frankfort, Fräulein Anny Weber in Burgebrach, and my colleagues Dr. Guttman and Dr. Mosbacher for their valuable help.

All investigations had to be made with the greatest care, as in no case could we divulge the source of the addresses or the reasons for the questions asked. In order to avoid prejudices I did not even inform my helpers of what I was really after. Whenever I suspected myself of being prejudiced, I endeavoured to bring in other helpers, whom I entrusted with clearing up the question of personal resemblance, without, however, informing them of the criminal records. In this way I think I did all I could to keep the investigation free from subjective influences.

In a large number of cases I have exact measurements; in almost all of them several photographs, in a lot of them duplicate observations as well as finger-prints. In three cases it happened that twins who had been summoned to my consulting-room for other alleged reasons told me spontaneously in the course of a lengthy conversation about their conflicts with the law. In other cases this did not happen, and even in asking general questions I met with strong resistance. One or two pairs resisted all attempts at investigation, and gave either the most

general or else unsatisfactory information. In other cases the question of monozygotism or dizygotism could not be cleared up satisfactorily owing to various difficulties. These cases I shall discuss later.

C. RESULTS

Thirty-seven pairs of twins were discovered and investigated in this way. In addition there were two other pairs, about whom I have a great deal of information, but whose derivation could not be ascertained.

The thirty-seven pairs included fifteen monozygotic and twenty-two dizygot couples.

In two cases of monozygotics and five of dizygotics neither twin had been imprisoned—these were pairs discovered among the brothers and sisters of ordinary prisoners. These have nothing to do with our question. Otherwise I should have to include as material the many pairs among my other investigations into twins who have never been imprisoned.

This leaves us with thirty pairs—thirteen monozygotic and seventeen dizygot, one of whom, i.e. the subject first investigated, had been imprisoned.

Among the thirteen monozygotic pairs the second twin was also imprisoned in ten cases, but in three cases had remained clear of the law. Among the seventeen dizygot pairs the second twin had also been imprisoned in two cases. In fifteen cases this had not occurred. *This leads us to the following conclusion: as far as crime is concerned, monozygotic twins on the whole react in a definitely similar manner, dizygot twins behave quite differently.* If, therefore, we attach importance to the twin method of investigation we must admit that

as far as the *causes of crime are concerned, innate tendencies play a preponderant part.*

One pair of dizygotic twins, both of whom were imprisoned, I should really place with the monozygotic. Two separate and reliable investigators took measurements which, together with photographs, showed complete agreement in unusual bodily characteristics, complexion, and so on. I am not including them because it is recorded that the twins were never taken for one another. This can be due to differences caused by environmental influences, to a different shape of the skull or the jaw-line, but I cannot be sure of it. I therefore prefer to leave them out.

I must just mention two pairs not included in the investigation. In one case they were twins of twenty-three years of age, one of whom stabbed the other to death. According to Par. 51 of the legal code the survivor was acquitted. No photographs of the victim exist. The relations are not intelligent, but they declare that there was a close resemblance between the two brothers. The record of the twin who was killed is apparently clean. But he was always threatening his twin, his other brothers and sisters, even his mother, with the knife, and like the survivor he was a bad lot.

Both brothers were feeble-minded, the survivor a little more than the dead one; both were hard of hearing and short-sighted. The living one was rather better-tempered and also more sentimental. Both of them had had many illnesses and had been badly knocked about by their drunken father. I saw the survivor and found there were still scars on his head. They were almost certainly monozygotic and should be classified with those whose conduct agreed.

The other pair were girls, now about twenty-four years old. Here are their records:

- | | |
|--------|---|
| ANNA. | 1. 1920. Lower Court. Unlicensed prostitution. Two weeks' detention. |
| | 2. 1922. Court of Jurors. Theft. Fourteen days' imprisonment. |
| | 3. 1924. Swindling. Four weeks' imprisonment. |
| | 4. 1924. Giving false information as to identity. Fourteen days' detention. |
| | |
| MARIA. | 1. 1921. Court of Jurors. Theft. Five days' imprisonment. |
| | 2. 1922. Court of Jurors. Theft. Three weeks' imprisonment. |

In this case the results of investigation were meagre. The twins themselves could not be got at. Their parents did not reply to the letters sent them. The records of the hospital where A. was taken were lost. The reformatories to which both were sent could give us no information with regard to our problem. We received helpful and detailed information from their birthplace, but the twins were there for a short time only. We do know that both were of low intelligence and were often ill, and that they were apparently not extremely alike. One of them had epileptic attacks. We have a clinical record of one of them, but not much can be got out of it. Finally, we know that the mother was a warehouse thief and systematically taught the children to steal. This fact makes them straightaway

unsuitable as material, in any case so long as they were still young. If classified at all on these facts they would have to be added to the monozygotics of similar behavior.

I just want to mention the brother and sister pairs about whom we accidentally got some information. In nine cases the twin brother was sentenced but not the twin sister. In one case the twin sister was imprisoned several times for theft and swindling and was finally given four years for fraud. The twin brother, who was business manager of a large hotel, once incurred a small fine during the inflation period for putting up his prices contrary to the law; he had raised the price of beer one day earlier than the law allowed. In other ways his conduct was without fault, and he felt that he had been unjustly condemned, as he claimed to have made a mistake. In the last case, one of triplets, the sister of sixteen years was very heavily punished with eight days' detention for staying away from school, while the two brothers had clean records.

If, therefore, we analysed these cases as unfavourably as possible from the

point of view of innate tendencies, the result would be as follows:

Among thirteen monozygotic pairs, ten would agree and three would disagree. Of thirty dizygotic pairs, five would agree and twenty-five would disagree.

If we distributed them according to what seems most highly probable, we should get the following proportions:

Of fifteen monozygotic pairs, twelve would agree and three would disagree; of twenty-eight dizygotic pairs, two would agree and twenty-six would disagree.

But these additions would not alter the basic proportions.

If we only took the numbers of which we are certain, the result would be that whereas 77 per cent. of monozygotic twins agree in their behaviour in relation to crime, only about 12 per cent. of the dizygotic twins do so. The proportions of these figures give some insight into the environmental influences, as does the lack of complete agreement of the figures regarding the monozygotics. But such numerical calculations have no real value in such complicated cases.

INTRODUCTION

Here is another selection taken from Watson's book, *Behaviorism*. In this chapter Watson is directly concerned with the relative contributions of innate and experiential factors in human development. At the time when Watson was writing, "instincts" were being invoked to explain an incredibly wide range of human behaviors. Towards the end of this selection, Watson mentions a list proposed by William James including shyness, modesty, kleptomania, and constructiveness as basic instinctual behavior patterns which human beings presumably inherit. In this chapter from his book Watson again argues from his firm behaviorist point of view: that while there are undoubtedly important inherited constraints on behavior (especially structural constraints due to changes in body makeup) the primary factors influencing behavior are experiential ones. As you read this article try to decide whether you think Watson is over-emphasizing the importance of "nurture."

ARE THERE ANY HUMAN INSTINCTS?

John B. Watson

When the array of facts about any subject is not very complete it is customary to announce a thesis, that is, state what one is going to try to prove and then try to prove it by a logical argument. We are all in that position today about man's equipment at birth. We have not the full complement of facts about the so-called "instinctive" nature of man; hence, please look upon the next few chapters both as logical presentations of what facts there are in the case and as a thesis which the behaviorist is trying to defend. I shall present the behaviorist's thesis first.

The Thesis Presented

Man is an animal born with certain definite types of structure. Having that kind of structure, he is forced to respond to stimuli at birth in certain ways (for example: breathing, heart beat, sneezing, and the like. A fairly full list will be given later on). This repertoire of responses is in general the same for each of us. Yet there

Reprinted with permission, from Behaviorism by John B. Watson, W. W. Norton Company, Inc. Copyright 1924, 1925 by Peoples Institute Publishing Co., Inc. Copyright renewed 1952 by John B. Watson. Revised edition Copyright 1930 by W. W. Norton & Company, Inc. Copyright renewed 1958 by John B. Watson, pp. 93-113.

exists a certain amount of variation in each—the variation is probably merely proportional to the variation there is in structure (including in structure, of course, chemical constitution). It is probably the same repertoire now that it was when the *genus homo* first appeared many millions of years ago. Let us call this group of reactions man's *unlearned behavior*.

In this relatively simple list of human responses there is none corresponding to what is called an "instinct" by present-day psychologists and biologists. There are then for us no instincts—we no longer need the term in psychology. Everything we have been in the habit of calling an "instinct" today is a result largely of training—belongs to man's *learned behavior*.

As a corollary from this we draw the conclusion that there is no such thing as an inheritance of *capacity, talent, temperament, mental constitution and characteristics*. These things again depend on training that goes on mainly in the cradle. The behaviorist would not say: "He inherits his father's capacity or talent for being a fine swordsman." He would say: "This child certainly has his father's slender build of body, the same type of eyes. His build is wonderfully like his father's. He, too, has the build of a

swordsman." And he would go on to say: "—and his father is very fond of him. He put a tiny sword into his hand when he was a year of age, and in all their walks he talks sword play, attack and defense, the code of dueling and the like." A certain type of structure, plus early training—*slanting*—accounts for adult performance.

The Argument in Its Defense

Let us start by saying that from now on man for us is a *whole animal*. *When he reacts he reacts with each and every part of his body*. Sometimes he reacts more strongly with one group of muscles and glands than with another. We then say he is doing something. We have named many of these acts such as breathing, sleeping, crawling, walking, running, fighting, crying. But please do not forget that each of these named acts involves the whole body.

We must begin, too, to think of man as a mammal—a primate—a two-legged animal with two arms and two delicate mobile hands. As an animal that has a nine-month embryonic life, a long helpless infancy, a slowly developing childhood, eight years of adolescence and a total life span of some three score years and ten.

We find this animal living in the tropics almost without shelter; going naked, living upon easily caught animals and upon fruit and herbs that require no cultivation. We find him in temperate regions, but dwelling here in well built, steam heated houses. We find the male always heavily clad even in summer, wearing a hat upon his head—the only naturally protected part of his body. We find the female of this species dressed in the scantiest of clothes. We find the male working frantically (the fe-

male rarely) at almost every kind of vocation, from digging holes in the ground, damming up water like beavers, to building tall buildings of steel and concrete. Again we find man in Arctic regions, clad in furs, eating fatty foods and living in houses built of snow and ice.

Everywhere we find man, we find him doing the strangest things, displaying the most divergent manners and customs. In Africa we find the blacks eating one another; in Southern China we find men eating rice and throwing it towards the mouth with dainty chopsticks. In other countries we find man using a metal knife and fork. So widely different is the adult behavior of the primitive Australian bushmen from that of the dwellers in internal China, and both of these groups differ so widely in behavior from the cultivated Englishman, that the question is forced upon us—*Do all members of the species homo wherever they are found in biological history start at birth with the same group of responses and are these responses aroused by the same set of stimuli?* Put in another way, is the *unlearned*, birth equipment of man, which we have been in the habit of calling *instincts*, the same wherever he is found, be it in Africa or in Boston, be it in the year six million B.C. or in 1930 A.D.? Has he the same unlearned equipment, whether born in the cotton fields of the South, in the Mayflower or beneath the silken purple quilts of European royalty?

The Genetic Psychologist's Answer

The genetic psychologist—the student best qualified to answer this question—hates to be faced with it because his data are limited. But since he is forced to answer he can give his

honest conviction. His answer is, "Yes, within the limits of individual variation, all men are born alike regardless of the station of their parents, regardless of the geological age in which they are born and regardless of the geographical zone in which they are born."

But one may say: "Is there nothing in heredity—is there nothing in eugenics—is there no advantage in being born an 'F. F. V.'—has there been no *progress* in human evolution?" Let us examine a few of the questions which excite so many people almost to the point of combat.

Certainly black parents will bear black children if the line is pure (except possibly once in a million years or so when a sport or "mutant" is born which theoretically may be white, yellow or red). Certainly the yellow skinned Chinese parents will bear a yellow skinned offspring. Certainly Caucasian parents will bear white children. But these differences are relatively slight. They are due among other things to differences in the amount and kind of pigments in the skin. It would be very difficult to take these infants at birth, study their behavior, and mark off differences in behavior that would characterize white from black and white or black from yellow. There will be differences in behavior but the burden of proof is upon the individual, be he biologist or eugenicist who claims that these racial differences are greater than the individual differences.

Again one hears the question: "How about children born from parents who have large hands, with short stiff fingers, extra fingers or toes? It can be shown that children from these parents inherit these peculiarities of structure." Our answer is: Yes, thousands of variations are laid down in

the germ plasm and will always appear (other factors being equal) in the offspring. Other inheritances are color of hair, color of eyes, texture of skin, Albinism (very light individuals with little or no pigment in hair and eyes—vision always being defective). The biologist, knowing the makeup of the parents and grandparents, can predict many of even the finer structural characteristics of the offspring.

So let us hasten to admit—yes, there are heritable differences in form, in structure. Some people are born with long slender fingers, with delicate throat structure; some are born tall, large, of prizefighter build; others with delicate skin and eye coloring. These differences are in the germ plasm and are handed down from parent to child. More questionable is the inheritance of such things as the early or late graying of hair, the early loss of hair, the span of life, the bearing of twins, and the like. Many of these questions have already been answered by biologists and many others are in the process of being answered. But do not let these undoubted facts of inheritance lead us astray as they have some of the biologists. The mere presence of these structures tells us not one thing about function. This has been the source of a great deal of confusion. Much of our structure laid down in heredity would never come to light, would never show in function, unless the organism were put in a certain environment, subjected to certain stimuli and forced to undergo training. Our hereditary structure lies ready to be shaped in a thousand different ways—the same structure—depending on the way in which the child is brought up. To convince oneself, measure the right arm of the blacksmith, look at the pictures of strong

men in our terrible magazines devoted to physical culture. Or turn to the poor bent back of the ancient bookkeeper. These people are structurally shaped (within limits) by the kinds of lives they lead.

Are "Mental" Traits Inherited?

But every one admits this about bone and tendons and muscles—"now how about mental traits? Does the behaviorist mean to say that great talent is not inherited? That criminal tendencies are not inherited? Surely we can prove that these things can be inherited." This was the older idea, the idea which grew up before we knew as much about what early shaping throughout infant life will do as we now know. The question is often put in specific form: "Look at the musicians who are sons of musicians; look at Wesley Smith, the son of the great economist, John Smith—surely a chip off the old block if ever there was one." The behaviorist recognizes no such things as mental traits, dispositions or tendencies. Hence, for him, there is no use in raising the question of the inheritance of talent in its old form.

Wesley Smith was thrown into an environment early in life that fairly reeked with economic, political and social questions. His attachment for his father was strong. The path he took was a very natural one. He went into that life for the same reason that your son becomes a lawyer, a doctor, or a politician. If the father is a shoemaker, a saloonkeeper, or a street cleaner—or is engaged in any other socially unrecognized occupation, the son does not follow so easily in the father's footsteps, but that is another story. Why did Wesley Smith succeed in reaching eminence when so many

sons who had famous fathers failed to attain equal eminence? Was it because this particular son inherited his father's talent? There may be a thousand reasons, not one of which lends any color to the view that Wesley Smith inherited the "talent" of his father. Suppose John Smith had had three sons who by hypothesis all had bodies so made up anatomically and physiologically that each could put on the same organization (habits) as the other two.¹ Suppose further that all three began to work upon economics at the age of six years. One was beloved by his father. He followed in his father's footsteps and due to his father's tutorship this son overtook and finally surpassed his father. Two years after the birth of Wesley, the second son was born; but the father was taken up with the elder son. The second son was beloved by the mother who now got less of her husband's time, so she devoted her time to the second son. The second son could not follow so closely in the footsteps of his father; he was influenced naturally by what his mother was doing. He early gave up his economic studies, entered society and ultimately became a "lounge lizard." The third son, born two years later, was unwanted. The father was taken up with the eldest son, the mother with the second son. The third son was also put to work upon economics, but receiving little parental care, he drifted daily towards the servants' quarters. An unscrupulous maid had taught him to masturbate at three. At twelve the chauffeur made a homosexual of him. Later falling in with neighborhood thieves he became a pickpocket, then a stool-pigeon and finally a drug fiend.

¹ And by this statement we do *not* mean that their genetic constitution is identical.

He died of paresis in an insane asylum. There was nothing wrong with the heredity of any one of these sons. All by hypothesis had equal chances at birth. All could have been the fathers of fine, healthy sons if their respective wives had been of good stock (except possibly the third son *after* he contracted syphilis).

Objectors will probably say that the behaviorist is flying in the face of the known facts of eugenics and experimental evolution—that the geneticists have proven that many of the behavior characteristics of the parents are handed down to the offspring—they will cite mathematical ability, musical ability, and many, many other types. Our reply is that the geneticists are working under the banner of the old “faculty” psychology. One need not give very much weight to any of their present conclusions. We no longer believe in faculties nor in any stereotyped patterns of behavior which go under the names of “talent” and inherited “capacities.”

Differences in Structure and Differences in Early Training Will Account for All Differences in Later Behavior

We have already asserted that even though there is individual variation in structure we can find no real proof that man's unlearned repertoire of acts has differed widely through the ages or that he has ever been either more or less capable of putting on complex training than he is in 1930. The fact that there are marked individual variations in structure among men has been known since biology began. But we have never sufficiently utilized it in analyzing man's behavior. I want to utilize another fact only recently brought out by the be-

haviorists and other students of animal psychology. Namely, that *habit formation starts in all probability in embryonic life and that even in the human young, environment shapes behavior so quickly that all of the older ideas about what types of behavior are inherited and what are learned break down*. Grant variations in structure at birth and rapid habit formation from birth, and you have a basis for explaining many of the so-called facts of inheritance of “mental” characteristics. Let us take up these two points:

(1) Human Beings Differ in the Way They Are Put Together

In our study of the makeup of the body, we obtained a faint idea of the complexity of the material and the processes that enter into the makeup of a human being. It enabled us to tolerate the fact too that there must be variation in the way these complicated tissues are put together. We have just brought out the fact that some human beings are born with long fingers, some with short; some with long arm and leg bones, some with short; some with hard bones, and some with soft; some with over-developed glands; some with poorly functioning glands. Again we know that we can identify human beings by differences in their finger-prints. No two human beings have ever had the same finger-prints, yet one can mark off man's hand and footprints from the tracks of all other animals. No two human beings have bones exactly alike, yet any good comparative anatomist can pick out a human bone (and there are over 200 of them) from the bones of any other mammal. If so simple a thing as the markings on the fingers differs in every individual, we

have absolute proof that general behavior will and must be different. Infants crawl differently, cry differently, differ in the frequency with which defecation and urination occur, differ in early vocal efforts, in requirements for food, in the speed and rapidity with which they use their hands—even identical twins show these differences—because they differ structurally and differ slightly in their chemical makeup. They differ, likewise in the finer details of sense organ equipment, in the details of brain and cord structure, in the heart and circulatory mechanisms and in the length, breadth, thickness and flexibility of the striped muscular systems.

Yet with all of these structural differences “a man’s a man for a’ that”—he is made up of the same material as other men and has the same general architectural plan regardless of habit.

(2) Differences in Early Training Make Man Still More Different

There are then admittedly these slight but significant differences in structure between each human being and every other human being. Differences in early training are even more marked. We now know that conditioned reflexes start in the human child at birth (and probably before)—we know that there is no such thing as giving two children, even though belonging to the same family, the same training. A doting young married couple have twins—a boy and a girl—the children are dressed alike and fed alike. But the father pets and fondles the girl, surrounds her with love; the mother treats the boy in the same way, but the father wants the boy to follow in his own footsteps. He is stern with him—he can’t help

shaping the boy his way. The mother wants the girl to be modest and maidenly. Soon these children show great differences in behavior. They receive different training from infancy. The next children are born. Now the father is more taken up with affairs—he has to work harder. The mother is more taken up with social duties; servants are brought in. The younger children have brothers and sisters; they are brought up in a world wholly different from that of the older children. One child falls ill. Strict training is abandoned, all rules are off with a sickly child. Again, one child gets badly frightened—becomes conditioned—shows fear at everything; he becomes timid and his regular course of boyish activity is interfered with. Indeed we may take an actual case. Two girls, aged nine, live in adjoining houses. They have the “same” training (mothers are close friends and bring up children according to the same rules). One day they took a walk. The girl on the left looked at the street and saw only street activity, the one on the right looked towards the houses and saw a man exposing his sex organs. The girl on the right was considerably troubled and disturbed and reached equanimity only after months of discussion with her parents.

The Conclusion We Draw

How will these two points explain the so-called facts of inheritance of talent or mental characteristics? Let us take a hypothetical case. Here are two boys, one aged 7, the other 6. The father is a pianist of great talent, the mother an artist working in oil, a portrait painter of note. The father has strong large hands but with long, flexible fingers (it is a myth that all

artists have long, tapering, finely formed fingers). The older son has the same type of hand. The father loves his first born, the mother the younger. Then the process of "creating them in his own image" begins. The world is brought up on the basis largely of shaping the young you are attached to as you yourself have been shaped. In this case the older becomes a wonderful pianist, the younger an indifferent artist. So much for different training or different slanting in youth. But what about different structure? Please note this. The younger son, under ordinary conditions, could not have been trained into a pianist. His fingers were not long enough and the muscular arrangement of the hand was not flexible enough. But even here we should be cautious—the piano is a standard instrument—a certain finger span and a certain hand, wrist and finger strength are needed. But suppose the father had been fond of the younger child and said, "I want him to be a pianist and I am going to try an experiment—his fingers are short—he'll never have a flexible hand, so I'll build him a piano. I'll make the keys narrow so that even with his short fingers his span will be sufficient, and I'll make a different leverage for the keys so that no particular strength or even flexibility will be needed." Who knows—the younger son under these conditions might have become the world's greatest pianist.

Such factors, especially those on the training side, have been wholly neglected in the study of inheritance. We have not the facts to build up statistics on the inheritance of special types of behavior, and until the facts have been brought out by the study of the human young, all data on the evolution of different forms of human behavior and eugenics must be accepted with the greatest possible caution.

Our conclusion, then, is that we have no real evidence of the inheritance of traits. I would feel perfectly confident in the ultimately favorable outcome of careful upbringing of a *healthy, well-formed baby* born of a long line of crooks, murderers and thieves, and prostitutes. Who has any evidence to the contrary? Many, many thousands of children yearly, born from moral households and steadfast parents become wayward, steal, become prostitutes, through one mishap or another of nurture. Many more thousands of sons and daughters of the wicked grow up to be wicked because they couldn't grow up any other way in such surroundings. But let one adopted child who has a bad ancestry go wrong and it is used as incontestable evidence for the inheritance of moral turpitude and criminal tendencies. As a matter of fact, there has not been a double handful of cases in the whole of our civilization of which records have been carefully enough kept for us to draw any such conclusions—mental testers, Lombroso, and all other students of criminality to the contrary notwithstanding. As a matter of fact adopted children are never brought up as one's own. One cannot use statistics gained from observations in charitable institutions and orphan asylums. All one needs to do to discount such statistics is to go there and work for a while, and I say this without trying to belittle the work of such organizations.

I should like to go one step further now and say, "Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select—doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants,

tendencies, abilities, vocations, and race of his ancestors." I am going beyond my facts and I admit it, but so have the advocates of the contrary and they have been doing it for many thousands of years. Please note that when this experiment is made I am to be allowed to specify the way the children are to be brought up and the type of world they have to live in.

Where there are structural defects that are inherited, as apparently is the case in certain glandular diseases; in "mental" defectives; where there is intra-uterine infection as in syphilis and in gonorrhea, troublesome behavior of one kind or another may develop early and rapidly. But some of these children haven't the structural possibilities to be trained—as when fundamental connections in body and brain are lacking. Again, where there are structural defects more easily observed as in deformities, loss of digits, possession of extra digits, there is social inferiority—competition on equal grounds is denied. The same is true when "inferior" races are brought up along with "superior" races. We have no sure evidence of inferiority in the negro race. Yet, educate a white child and a negro child in the same school—bring them up in the same family (theoretically without difference) and when society begins to exert its crushing might, the negro cannot compete.²

The truth is society does not like to face facts. Pride of race has been

strong, hence our Mayflower ancestry—our Daughters of the Revolution. We like to boast of our ancestry. It sets us apart. We like to think that it takes three generations to make a gentleman (sometimes a lot longer!) and that we have more than three behind us. Again, on the other hand, the belief in the inheritance of tendencies and traits saves us from blame in the training of our young. The mother says when her son goes "wrong"—"Look at his father" or "Look at his grandfather" (whichever one she hates). "What could you expect with that ancestry on his father's side?" And the father, when the girl shows wayward tendencies—"What can you expect? Her mother has always let every man she came in contact with make love to her." "If these tendencies are inherited we can't be much blamed for it. Traits in the older psychologies are God-given and if my boy or girl goes wrong, I as a parent can't be blamed."

During the past five years considerable light has been thrown upon the relative influence of environment and heredity by the study of identical twins. Let us examine this material at some length.

Does the Experimental Work on Identical Twins Negate Our Emphasis on the Relatively Greater Influence of Environment in the Early Differentiation of Behavior?

Dr. Arnold Gesell is the strongest advocate of heredity and growth factors. In his article, "Behavior Resemblance in Identical Infant Twins" (*Eugenical News*, Vol. XIV, No. 5, May, 1929) he states very positively that the behavior pattern of monozygotic twins shows a special degree of resemblance in behavior. For example, "The postural, perceptual,

² I say nothing here on the inheritance of acquired behavior characteristics. The evidence in biology is all against it. An infant descended from a long line of blacksmith ancestors starts with approximately the same puny upper right arm girth as his original blacksmith forebear—and with one no larger than his own left arm. The facts for and against have been beautifully summarized in Jennings' *The Biological Basis of Human Nature*, page 338 ff.

prehensory, exploitive and imitative responses to one, two and three one-inch red cubes placed before each child were described and photographed to provide objective data for a detailed study of the development of behavior patterns." Twins T and C under these test conditions showed a marked degree of resemblance in behavior patterns. These correspondences in behavior patterning were literally uncountable. For example, there were 99 items of minor disparity and 513 items of identical or nearly identical correspondence.

While these experiments are interesting, they constitute no proof that identity of patterns in these two youngsters was any closer than would have been the case had he investigated infants of the same age and kept under relatively the same environmental conditions, namely residence in the clinic, provided the children had been of the same general weight and physical makeup and previous conditioning. So far as I have seen from the report, no controls were introduced in the experiment. Dr. Gesell concludes: "During the course of our simultaneous observation of the behavior patterns of Twins T and C, the correspondence frequently was so striking as to recall Galton's metaphor about clocks and necessitarian whirring wheels!"

That Dr. Gesell is not supported in this conclusion by any other similar study appears from the work of Muller and of Newman. H. G. Muller, *Journal of Heredity*, December 1925, describes his tests upon identical twins, B and J. They were 30 years old at the time of the test. They were separated at two weeks of age. One (B) lived in Wyoming at the time but had lived in New York, Washington and other places, and the other in Arizona (J).

B was brought up by foster parents who did mining, logging and hauling. As a child B was out with the teams all day and had only four years of formal schooling all told, including nine months in a business college. At 15 B obtained a clerical job. Her business career was not very successful—she tried in addition to clerical work, administrative and secretarial work.

J was also brought up by foster parents who owned a ranch and a roadhouse where things were lively. She was also out of doors a great deal and like B was a tomboy. She went through school including high school and then had some summer university work. She taught school and had a child.

Neither of the twins had been very well and their illnesses had been quite similar.

Their scores on the general intelligence tests (Otis and Army Alpha) were not very different but on some sections J was good where B was poor and *vice versa*. In other tests, for example, Kent Rosanoff's Association Test, J, on account of greater schooling responded almost twice as fast. On the tapping test, B because of the typewriting experience had a score of 207 against J's 164 in the same period of time. In the crossing test B had a score of 63 to J's 55. Muller concluded that all tests on motor reaction time, "will temperament" emotions and social attitudes "give results in striking contrast with those of the intelligence tests, in that the twins made markedly different scores on all these tests. *The differences* were, on the average, slightly greater than the median differences between the scores of two individuals chosen at random from the group in which the 'norms' of the tests had been established."

Newman in his article, "Mental

Traits of Identical Twins Reared Apart," *Journal of Heredity*, April, 1929, gives an even more striking case of identical twin boys reared apart—one in the city, the other in the country. Twins C and O were born in 1925. They are almost indistinguishable in physical appearance. I quote him: "In this case the environment and training of the twins were closely similar in general, at least insofar as we were able to ascertain the facts. Yet the 'personalities' of the two boys are utterly different. So opposite is the impression made by them upon all the observers that we were forcibly struck by it from the beginning."

On the Pressey test of emotional reaction, C's (city boy's) total score was 29; O's (country boy's) was 55, O thus being decidedly more neurotic. On the international group intelligence test, C's score was 156, while O's was 146. On Thurstone's Psychological Examination, C's score was 101 and O's 84. This test involves completion, artificial language, analogies, arithmetic, and opposites.

Newman studied also two pairs of identical twin sisters. The first pair, separated at the end of 18 months, were tested after 18 years (they had lived together one year before the psychological tests were made). During the separation the environments were very different. "The twins dealt with in this paper differed tremendously in mental capacity—the difference being three times as great as the average of 50 pairs of identical twins reared together." In this case the twins showed considerable resemblance in emotional characteristics (which was to be expected, due to the fact that the first 18 months of their lives were spent together—so early are emotional patterns laid down).

In the case of the second pair of identical twin girls studied by New-

man, the environments were also different. One attended school seven years longer than the other. They also were separated after living together 18 months. They were tested after 19 years of separation. The more highly educated twin, G, gave a very superior performance in every test of "mental" ability and "native" performance. In emotional tests they too were quite similar.

In all these tests on identical twins, I see not one scintilla of evidence for the contention of the biologist that the number and the relation of the genes are the determining factors in behavior patterns. In identical twins we are dealing with the same set of genes, and yet the outcome of different training yields different individuals.

Just to emphasize our point here again—after all, life outside the laboratory has a good deal of similarity. Suppose now we were to take identical twins into the laboratory and begin rigidly to condition them from birth to the twentieth year along utterly different lines as diverse as possible. We might even condition one of the children to grow up without language. Those of us who have spent years in the conditioning of children and animals cannot help but realize that the two end products would be as different as day is from night.

Yet in face of this evidence, Jennings, who is by all odds the fairest and most scientific of the biologists and most interested in genetics, in "*The Biological Basis of Human Nature*" says:

"Thus the study of the four cases of identical twins reared separately is favorable to the claims of those who hold that environment and experience may have a great effect on mental and on temperamental characteristics, producing marked difference in these re-

spects, even in individuals that are genetically alike. The four cases, however, support equally the view that the genetic constitution likewise affects greatly these matters. For in all cases the twins reared under diverse environments were much more alike in certain respects—in some cases in mental processes, in others in temperament—than can be accounted for on any basis except their identity in genetic constitution. The results of the study agree thoroughly with the conclusion, practically certain on other grounds, that both the genetic constitution and the environment deeply influence mental and temperamental characteristics; and that effects produced in one case by genetic constitution may be produced in another case by environment."

I am sure that Jennings, if he were to take account of the fact that in several of the above cases the children had had their early conditioning together, that the tests were by no means sufficient to bring out all or even a large part of the actual differences in behavior patterns, and that in the study of these cases the behaviorist had had no opportunity to train the children experimentally along different lines, would have to agree that my position, which I have outlined above, is the only one which has any reasonable facts and hypotheses to support it. And we must remember here that in identical twins we are taking the biologically hardest case for the behaviorist to cope with in his emphasis on environment.

Let us, then, lay aside the thought of inheritance of aptitudes, of "mental" characteristics, of special abilities (not based upon favorable structure such as throat formation in singing, hand formation in playing, structurally sound eyes and ears) and take up the more general question of what

the world has been in the habit of calling instincts.

Are There Any Instincts?

It is not easy to answer this question. Up to the advent of the behaviorist man was supposed to be a creature of many complicated instincts. A group of older writers, under the sway of the newly created theories of Darwin, vied with one another in finding new and perfect instincts in both man and animals. William James made a careful selection from among these asserted instincts and gave man the following list: *Climbing, imitation, emulation, rivalry, pugnacity, anger, resentment, sympathy, hunting, fear, appropriation, acquisitiveness, kleptomania, constructiveness, play, curiosity, sociability, shyness, cleanliness, modesty, shame, love, jealousy, parental love.* James states that no other mammal, not even the monkey, can lay claim to so large a list.

The behaviorist finds himself wholly unable to agree with James and the other psychologists who claim that man has unlearned activities of these complicated kinds. We have all been brought up on James or possibly even on a worse diet, and it is hard to run counter to him. James says an instinct is "a tendency to act in such a way as to bring about certain ends without having foresight of those ends." Surely this formulation fits a lot of the early behavior of children and young animals. At first this formulation looks convincing. But when one tests it out in terms of one's own observations on young animals and children, one finds that one has not a scientific definition but a metaphysical assumption. One gets lost in the sophistry of "foresight" and "end."

No subject in psychology today is

more written about than the so-called instincts. In the past few years many hundreds of articles have been written about instincts. The articles in general are of the armchair variety written by men who have never watched the whole life history of animals and the early childhood of the human young. Philosophy will never answer any questions about instincts. The questions asked are factual ones—to be answered only by genetic observation. Let me hasten to add that the behaviorist's knowledge of instinct also suffers from lack of observed facts but one cannot accuse him of going beyond natural science in his inferences. Before attempting to answer the question "What is an instinct?" let us take a little journey into mechanics. Possibly we may find that we do not need the term after all.

A Lesson from the Boomerang

I have in my hand a hardwood stick. If I throw it forward and upward it goes a certain distance and drops to the ground. I retrieve the stick, put it in hot water, bend it at a certain angle, throw it out again—it goes outward, revolving as it goes for a short distance, turns to the right and then drops down. Again I retrieve the stick, reshape it slightly and make its edges convex. I call it a boomerang. Again I throw it upward and outward. Again it goes forward revolving as it goes. Suddenly it turns, comes back and gracefully and kindly falls at my feet. It is still a stick, still made of the same material, but it has been shaped differently. *Has the boomerang an instinct to return to the hand of the thrower?* No? Well, why does it return? Because it is made in such a way that when it is thrown upward and outward with a given force it

must return (parallelogram of forces). Let me call attention to the fact here that all well made and well thrown boomerangs will return to or near to the thrower's feet, but no two will follow exactly the same forward pathway or the same return pathway, even if shot mechanically with the same application of force and at the same elevation, yet they are called boomerangs. This example may be a little unusual. Let us take one a little easier. Most of us have rolled dice now and again. Take a die, load it in a certain way, roll it, and the face bearing "six" will nearly always come up when the die is thrown. Why? The die must roll that way because of the way it was constructed. Again take a toy soldier. Mount it on a semi-circular loaded rubber base. No matter how one throws this soldier, he will always bob upright, oscillate a bit, then come to a steady vertical position. *Has the rubber soldier an instinct to stand erect?*

Notice that not until the boomerang, the toy soldier and the die are hurled into space do they exhibit their peculiarities of motion. Change their form or their structure, or alter greatly the material out of which they are made (make them of iron instead of wood or rubber), and their characteristic motion may markedly change. But man is made up of certain kinds of material—put together in certain ways. If he is hurled into action (as a result of stimulation) may he not exhibit movement (in advance of training) just as peculiar as (but no more mysterious than) that of the boomerang?³

³ One may argue that in mechanics action and reaction are equal—that the thrower imparts to the boomerang a quantity of energy equal to so many dynes and that just that same quantity of energy is used up by the

Concept of Instinct No Longer Needed in Psychology

This brings us now to our central thought. If the boomerang has no instinct (aptitude, capacity, tendency, trait) to return to the hand of the thrower; if we need no mysterious way of accounting for the motion of the boomerang, if the laws of physics will account for its motions—cannot

boomerang in returning to the thrower (including the heat loss to the air). When I touch a man with a hair, though, and he jumps two feet, the reaction is out of all proportion to the energy in the stimulus. The explanation is that in man the energy used in the reaction was stored. In dynamics you find the same thing when a match touches off a powder blast or when a breeze blows a rocking boulder from a cliff and the boulder destroys a house in the valley.

psychology see in this a much needed lesson in simplicity? Can it not dispense with instincts? Can we not say that man is built of certain materials put together in certain complex ways, and *as a corollary of the way he is put together and of the material out of which he is made—he must act (until learning has reshaped him) as he does act?*

But one may argue: "That gives your whole argument away—you admit he does a lot of things at birth which he is forced to do by his structure—this is just what I mean by *instinct*." My answer is that we must now go to the facts. We can no longer postpone a visit to the nursery. I think the student will find there, in the study of the infant and child, *little that will encourage him to keep sacred James' list of instincts.*

INTRODUCTION

This selection is an entire chapter from the book *Twins and Supertwins* by Amram Scheinfeld. This book was written, among other reasons, to tell the parents of twins what to expect as their children develop. Even if you don't happen to have any twin children, you can learn a good deal from this very readable chapter about the relative contributions of nature and nurture in the development of identical twins. Nowhere do we have a more promising source of data for sorting out the highly interrelated contributions which these two factors make.

IDENTICAL TWINS: MENTAL TRAITS

A. Scheinfeld

Joey and Jimmy, age $3\frac{1}{2}$, are typical peas-in-a-pod look-alikes. But their worried mother writes me, "At the nursery school I was told that Joey has an IQ of 112, Jimmy only 106. How can this be if they're *identical* twins?"

As we saw in earlier chapters, it is the comparison between twins, especially with respect to mental traits, that worries many parents. Suppose,

Reprinted from Twins and Supertwins by Amram Scheinfeld. Copyright © 1967 by Amram Scheinfeld. Reprinted by permission of J. B. Lippincott Company.

then, we do make comparisons: If *you* have identical twins, should you expect their minds, school achievements, and personalities to be alike? And if not, what can cause the differences, and how far can these go?

Since identical twins have exactly the same hereditary equipment for the same hereditary equipment for the mental as well as all other traits, the question goes beyond their relationship and into a much wider area: To what extent are the mental traits of human beings in general produced by their hereditary makeup, and to what extent by their rearing, experience

and education? We will find no easy answers. Despite decades of study by psychologists, there is agreement only that whatever the *inborn* mental capacities of people may be, their mental performance—including their scores on intelligence tests—may be much influenced by their environments and training. Most interesting for us, some of the best studies on the “heredity vs. environment” question have centered on identical and fraternal twins, in comparison with each other and with single siblings.

The IQs of identical twins of a pair usually tend to be very close together, although seldom exactly alike. In most cases their IQ difference is small—averaging little more than 5 points. This is considerably less than the usual IQ difference between members of a fraternal pair, which averages about 8.5 points, and is much smaller than the average difference between any two single children in a family, which may be from 10 to 14 points.

What are we to make of these figures? Are the IQs of identical twins so much more alike as a rule than those of two fraternal twins or two single siblings because they have identical hereditary factors and the others do not? Or is it because identical twins are reared more closely together and in the same way?

A first thing to note is that fraternal twins, too, are usually closer in their IQs than are two singletons in a family, even though the fraternal twins may have just as much difference in their “mental” genes. This shows that the similar environment and training of twins do help to make their minds work more like. It might be argued, then, that identical twins are most alike in IQ because they are most alike in training. So let’s ask: “Suppose a pair of identical twins (yours,

for example), were *not* reared together—that they were completely separated, not only in school, but were also raised in different homes? How would their IQs compare under those circumstances?”

The question has an answer in the cases of the separated twins we discussed [previously]. Tests of these identicals who had been reared apart from infancy onward showed that their IQs differed on the average by only a few points more than those of identicals who had been reared together. Drs. L. Erlenmeyer-Kimling and Lissy F. Jarvik have gone over the combined studies of IQs of thousands of twins and singletons, and have analyzed the “correlations” (statistical relationships) between the IQs of identicals reared together and apart, as well as of fraternal twins, pairs of single siblings, and pairs of unrelated children brought up together in foster homes or institutions.* The results shown in Illustration, p. 56, were considered evidence that the minds of identical twins do tend to be geared by their matching genes to perform in much the same way, even under different conditions.

A striking case in point was that of

* “Correlations” are degrees of closeness as measured by certain standards. If in a group of paired individuals—twins or others—the IQ of one member of a pair is always matched exactly by that of the other, this is a correlation of “100” (or “+1.00”). If no consistent relationship at all is found between the IQs of paired members in the group (for instance, if in one pair one member has an IQ of 110, the other of 90; in the next pair, one an IQ of 85, the other of 130, and so on) this would be a correlation of “zero.” In between the 100 and zero extremes, the more or less often the IQs of members of pairs match each other or come close together, the higher or the lower would be the average correlation for the group as a whole.

the identical British twins, sons of an Oxford tutor who had died just before they were born. The mother could keep only one, and so gave the other for adoption. The twin George, who remained with his mother, had a brilliant academic career. One day a former teacher of his, seeing a man whom he assumed to be George, greeted him heartily. Surprisingly, there came a puzzled answer in a strong, uncultured Welsh accent. It turned out, of course, that the man was the other twin, Llewellyn, who had been reared by his foster parents on an isolated farm in Wales. Yet when tested, while he was found very backward in reading and verbal performance, his IQ was 137—fully up to that of his educated twin, George.

There have been instances, however, in which the matching genes of separated identicals did not rule the mental roost. In one case there was a difference of 24 points in the IQs of two young women identicals separately adopted as infants and raised under very different conditions. Gladys, brought up in an isolated Canadian mountain district and with a back-country grade-school education, scored only 92 in IQ. Her twin, Helen, college-educated and a teacher, scored 116. But this was the most extreme case discovered. At that, the IQ difference—quite likely due to the markedly superior education, and training in taking tests, of the one twin—was not nearly so great as can often be found between fraternal twins or singletons raised together in a family.

In another well-known study, two Yale experts, the late Dr. Arnold Gesell, and Dr. Helen Thompson, sought to learn what would happen if two highly similar three-and-one-half-year-old girl identicals were con-

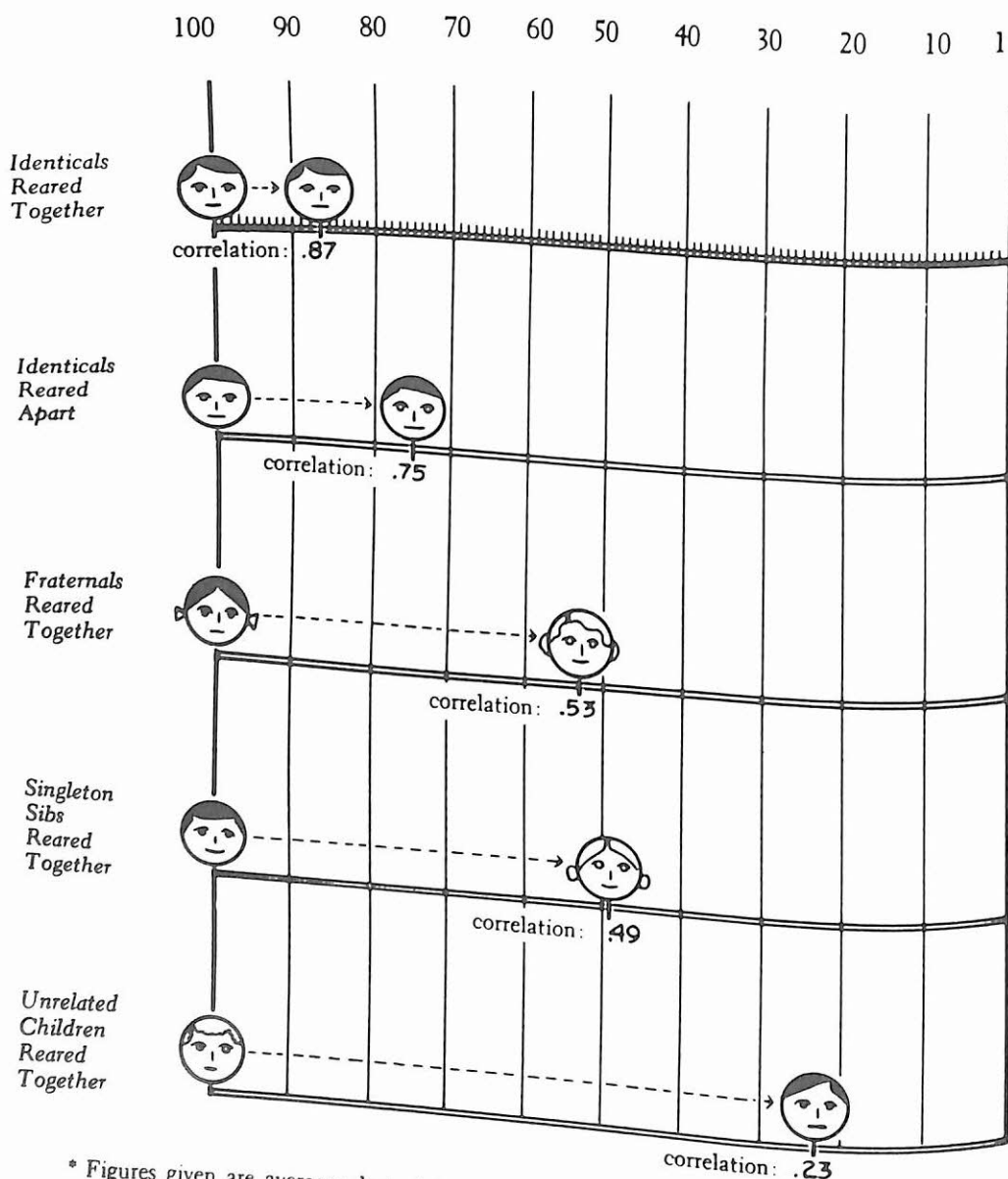
ditioned in very different ways. Accordingly one of the twins was given intensive special training for a long period, while the other was not. But when the separate training of the one twin ceased, and both were left to their own devices, it was not long before they bounced back together in performance. So it would appear that *under the same conditions* the minds of identical twins can be expected to work in about the same way; and it is only under extraordinary circumstances that they will show great difference in mental performance.

Yet even small differences between identical twins as they grow up may perplex and worry parents of the twins concerned. The first thing to explain, then, is that intelligence tests or personal observations of young children may be very unreliable (and under age 2 are often meaningless). Except where there are clearly-established mental defects, one may have to wait until twins are about six years old before their IQs can be considered indicative of their future mental performance, either as a pair or in relation to each other. This is especially so if the twins were premature, and most certainly true if one was much behind the other in development at birth. In most cases where both twins seem somewhat backward in early childhood (but are not defective) they may be expected to achieve average IQs later. Similarly, when one twin has been behind the other in IQ, he may catch up before long and sometimes—like the tortoise and the hare—come out ahead.

Also to be considered is the fact that the IQ lumps together the scores on various types of mental performance—language, numbers, abstract reasoning, and so on. While one twin may have a total IQ score higher than the

HOW CLOSE ARE TWINS' IQs?

Here are the IQ correlations (degrees of closeness) between pairs of twins of different kinds, as compared with those of other paired children.*



* Figures given are averages derived from an analysis of intelligence test scores in 52 studies, covering thousands of twins and nontwin children, as reported by Drs. L. Erlenmeyer Kimling and Lissy F. Jarvik (*Science*, 142, Dec. 12, 1963). "Unrelated children" refers to foster children of different parents reared in the same home. "Correlation" is further explained in the Text.

other, his twin may do better on some parts of the test. The twin who talks more fluently may not be as good at working out problems as his twin and vice versa. Or, again, both twins may end up with almost the same IQ totals, but may have gotten their scores in different ways. Thus, it has been found that while the IQs of identical twins may be very similar, their school achievements may be less similar; and the more the twins have been separated, the less similarity there is apt to be in their scores in such subjects as reading, spelling and arithmetic. However, there does appear to be a strong hereditary component in *mathematical ability*, and in *musical talent*, and perhaps also in aptitude for *science*; so, if two identical twins differ greatly in achievement in these areas, there may be special reason to assume that one is lagging behind his capacities.

Now suppose there is an unmistakable and large disparity in the whole intelligence of one identical twin as compared with the other. What can account for it? Sometimes this may be traced to physical factors—perhaps to some injury or deficiency which the lagging twin had suffered before or during birth, or to some disease or mishap in infancy. Some—but not all—types of epilepsy and cerebral palsy may cause mental retardation, and may occur in one identical twin and not in the other. Also, in former years, excess oxygen in incubators may have affected the intelligence (as well as the eyes) of many prematurely-born twins, and perhaps one more than the other. More often, any significant lag on the part of one identical twin either in IQ or in a particular subject can be traced to psychological factors. Remembering that Jimmy has in all probability the same basic mental equipment as Joey, the reasons he's

not using it as fully may be much like those which inhibit any child from doing as well as he might: shyness when called on to recite; lack of drive or interest; a fancied resentment. In many cases alert parents and/or teachers may be able to detect or guess what's wrong. If not, and should a pediatrician find nothing physically wrong with the lagging twin, it may be well to consult a psychologist either in the school system or outside.

Relationships between the twins can affect their relative mental performances and should not be overlooked. One factor may be the frequently observed tendency among twins for one to be dominant, and the other to take a backseat. Sometimes this may lead to the dominant twin being more aggressive, enterprising and successful in school, and more popular with teachers and classmates. But not infrequently the twin who is less successful in social relationships or in physical feats may seek an outlet in his studies and do better in his schoolwork. This type of compensation may tend to keep the twins in balance with each other—which is just what their parents may wish most.

In fact, it is not necessary that identical twins should be equally successful in everything. So long as each is helped to develop his mental capacities as fully as he wishes, and in a way to bring him happiness, there is no reason why differences in school performance or in various competitive areas should keep twins from being well-adjusted to each other and later becoming well-adjusted adults. Nevertheless, while up to a point differences in identical twins in given subjects, aptitudes or interests may be encouraged as leading to individuality development, there is danger if they do not keep fairly abreast of each other in their total school averages and relative

achievements. If one twin consistently falls behind the other to any marked degree, the situation may become serious with the approach of maturity and/or the college years. One twin may miss out on entrance to a preferred college, or fail to have equal opportunity for getting a job or attaining success and happiness. So, as a preventive measure, beginning with the twins' early school years, parents should make every effort to see that the achievements of one twin—scholastic, social, athletic—come up to and average out fairly closely to those of the other.

This brings us to the next question: How alike can one expect identical twins to be in *personality*? Often when this is asked we find an interesting psychological quirk—not in the twins, but in their parents. For, in many cases, when comparisons are resented with respect to school performances, and mental differences are deplored, parents don't seem to feel the same way about comparing twins' personalities. Indeed, many mothers, either because of desire to distinguish between their identical twins or to secure assurance that the twins' individualities are developing, seem eager to detect differences in personality as soon as possible. Thus, one of the Mothers of Twins clubs routinely queried enrolling members—many of whose twins were mere infants—as to which twin was “dominant” or “submissive.” Among the many mothers reporting such a distinction was one whose twin boys were *seven months old*. (One wonders in what sort of conflict—perhaps a “bottle battle”?—these teensie-weensies had settled the question of supremacy.) Also, in statements to me, a mother of sixteen-month-old girl twins labeled one an extrovert, the other an introvert, the

mother of eight-month-old boy twins said one was a real “he-man,” the other “more of the reflective and delicate type”; and various other parents have professed to see in identical twins who were anywhere from four months to a year old, such personality contrasts as “bravery” versus “timidity,” “sociability” versus “aloofness,” “flirtiness” versus “prissiness”—and indeed almost any contrasting traits one can think of.

The fact is, of course, that professional psychologists would hesitate to make such definite decisions about the personalities of any two very young children, for personality tests in the preschool years may be even less reliable than are the intelligence tests. To be sure, at any given time certain differences in personality can be detected in young identicals, but these are usually minor and apparent mainly to their parents, siblings or intimate friends. Moreover, many parents report the familiar seesaw effect of a given trait shifting from one twin to the other at various times, and quite often it is merely an illusion that the contrasts exist at all.

When identical twins do reach the age when their personalities can be judged with any degree of reliability, the findings of many studies show this: In almost all basic personality traits identical twins tend to be much more alike than do fraternal twins of the same sex, or two single siblings in a family. For example, in one study of twins at ages 11 to 15, by Drs. Raymond B. Cattell, Duncan B. Blewett, and John R. Beloff, such traits as being easygoing, sociable and warm-hearted or bold and outgoing were shown as appearing together in both members of an identical pair much more often than in both twins of a fraternal pair or in two single siblings.

In another study, Dr. Steven G. Vandenberg (University of Louisville) found much more similarity between the identicals than between the fraternal twins in such traits as response to music and color, and in degrees of self-confidence and stubbornness.

Food preferences are of special interest in identicals. As a rule, both will go for or reject much the same foods and beverages, sometimes from their earliest years. In the case of my identical-twin doctor friends, told about in the previous chapter, it was fascinating to hear them discuss their likes and dislikes. Throughout their lives they had found that if one didn't like something, the other usually didn't, either. For instance, both had discovered separately that they disliked peppermint, grapefruit, tomato soup, fruit juices, and cocktails.

That there may be an inborn basis for food and drink preferences or aversions in identical twins, arising from their biochemical functioning, has been previously suggested. Again, we noted that allergic reactions of identicals to various specific substances may be much the same. Among other things, biochemist Roger J. Williams has maintained that a craving for alcohol by some individuals, or a dislike for it or inability to tolerate it by others, may be linked with inherent characteristics of the body chemistry. If so, there would be a special reason why the drinking habits or aversions of identical twins should tend to be the same. Similarities in the smoking habits of identical twins—indulgence or abstinence—were also reported by the late British geneticist, Sir Ronald Fisher, as pointing to hereditary influences.

But the question always pops up: "Doesn't the fact that the *environments* of identical twins are so much

the same account for their similarities in personality?" To a considerable extent this may be so, as it applies to their early closeness and training. But it is important to note that identical twins themselves *create* similarities in their "personality environment" by looking so much alike and functioning alike in so many ways. In other words, whether they are both good-looking or homely, robust or sickly, bright or dull, talented or unartistic, identical twins tend to inspire much the same reactions from other people, and this in itself is a great influence in molding their personalities along similar lines.

Thus, it is again not surprising that even among the twins who are separated in infancy and reared apart, close resemblances in personality often develop. One of the most striking cases on record was that of the twins Millan and George, born in Salt Lake City, who had been separated almost immediately after birth and brought up far apart by foster parents of different types, under quite different conditions. When they came together, at age 19, they were found to be remarkably alike not only in looks but also in intelligence, temperament, and achievement. Both had artistic leanings, one toward music, the other toward drawing; and both had won amateur boxing championships. (Their twinship further asserted itself in a sad way when, while serving apart in the Army, they both developed—almost simultaneously—the same rare crippling disease of *spondylitis*.) Also, there were the Baylor University twins, Louise and Lola, whose amazing physical similarity after years of separation was mentioned in a previous chapter. These twins, too, on being reunited as young women, were found to be closely matched not only

in IQs but also in many abilities (both were artistic) and in personality traits.

Added to the accounts of American reared-apart twins is that of the two male identicals given the names of Tony Milasi and Roger Brooks after their separation in infancy and adoption by two quite different couples, one Italian, the other Jewish. With Tony raised in New York State and Roger in Florida, these twins knew nothing of each other's existence until, through the now familiar coincidence of mistaken identity, they finally were brought together at age 23 (in 1964). Though psychological tests by Dr. Syvil Marquit of Miami showed that they had nearly the same IQs and basic aptitudes, their personalities were apparently very different (not surprising in view of how different had been their home lives and rearing), Roger being rated as "more sensitive and impressionable," Tony as "more extroverted and self-assured." (Nonetheless, the twins were so drawn together that Roger moved to Tony's city to live with him.)

Other studies of identical twins reared apart, made in Denmark by Dr. N. Jule-Nielsen, give further evidence of how matching hereditary factors can lead to the development of great similarities in twins' personalities, even when there have been great environmental differences. The Danish group comprised twelve pairs of adult identical twins, ranging in age from twenty-two to seventy-seven, who had been separated in early childhood and not reunited until much later—in many cases, as with twins previously mentioned, because one was repeatedly mistaken for the other. One pair, the male identicals Palle and Peter, had been raised from infancy in very different homes in Copenhagen, and had quite marked differences in edu-

cation (one had only an elementary schooling, while the other was studying medicine); but, when brought together at age 22, they proved to be remarkably similar in mental development and personality. As Dr. Jule-Nielsen reported: "Both seemed introspective, lacked self-confidence, found it difficult to achieve emotional contact with other persons, and presented similar neurotic, hypochondriacal symptoms." Despite this, they took to each other so well that they left their respective homes and thereafter kept closely together.

Not so successful was the reunion of another identical male twin pair, Kaj and Robert, who had been apart from infancy until the age of 40. Kaj, raised in a village by a poor couple, had a jail record and an erratic history. Robert, who grew up and was educated under favorable conditions, had become a technician. But he too, though never in direct conflict with the law, had, like his twin, showed an antisocial streak, was maladjusted, and had sought psychiatric help. Tests showed them to be much alike in basic personality and with much the same problems—work, money and women. (Each had been married three times.) Before long, they broke off relations with each other.

Also unsuccessful was the reunion of two women identicals, Ingrid and Olga, who had been separated when only eight weeks old. Ingrid was reared on a farm in one part of Jutland, Olga in a children's home in another part. When brought together again at age 35, they found that although they had similar personalities, they were too different in background and views to provide any basis for a continuing close relationship.

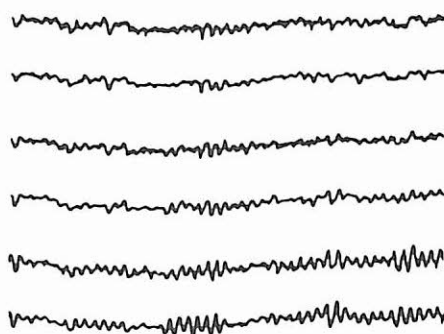
Of special interest are the brain-wave recordings (electroencephalo-

grams) of the group of separated Danish twins which were made by Dr. N. Jule-Nielsen in association with Dr. Bent Harvald. The "EEGs" of the two members of each pair were found to be almost completely alike in all significant aspects (as are those

similarly despite differences in environment. But one could also give many instances in which identical twins, separated or not, showed considerable differences in personality (as was true with Tony and Roger, previously mentioned). In a later chapter

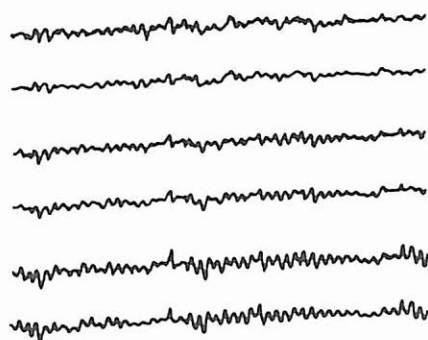
BRAIN WAVES OF SEPARATED IDENTICAL TWINS

INGRID'S EEGS



150 μ V

OLGA'S EEGS



The above amazingly similar brain-wave patterns (encephalograms, or "EEGs") are those of the two thirty-seven-year-old Danish identical twins, Ingrid and Olga (referred to in the text), who had been separated at the age of eight weeks, reared apart under very different conditions, and not reunited until about a year before the recordings were made by Drs. N. Jule-Nielsen and Bent Harvald.

of Ingrid and Olga, which are reproduced here). This led to the conclusion that brain waves in individuals generally are primarily determined by hereditary factors which continue to assert themselves throughout life.

We do not wish to leave readers with the impression that the personality patterns of identical twins (or of single individuals, for that matter) are unalterably set for life by their genes, and that environmental influences play only a small part. The cases usually dwelt on (because they may be the most interesting) are mainly those of twins whose personalities developed

we will tell of some extreme cases, in which one twin became a criminal or was a suicide, or the behavior paths diverged sharply in some other respect. Perhaps, just as parents may look for and emphasize very small differences between their identical twins, many scientists in studying twins may have been inclined to look for and emphasize similarities.

Regretfully for scientific study but happily for twins, the cases of those separated from infancy onward are becoming increasingly rare. The sharp decline in parental mortality has greatly reduced the chance of twins

being orphaned in infancy; public aid has enabled impoverished parents to keep their twins; and also, when twins have had to be given out for adoption, the demand for these babies (at least in the United States, and in other advanced countries) has become so great that any twins offered are eagerly snapped up as a pair. (Among the 500 or so pairs of twins reported on in the questionnaires filled out for me, at least six pairs, to my knowledge, were adopted. In another case, reported to me, a Los Angeles couple, in 1960, adopted a set of triplets.) However, the stories already available of twins separated in the past, and other evidence brought out in this chapter, should help to give parents a good idea of what can be looked for in the way of mental and personality differences or likenesses in their own identicals, and what practical value this knowledge can have.

Above all, one should keep in mind that identical twins are in a unique position in relation to themselves, and in the way other people react to them. Comparisons between the twins—no matter what the parents may feel or do—are unavoidable. These will be made not only by other people but also by the twins themselves throughout their lives. If comparisons merely reveal interesting differences between the twins, this may be all to the good. Harm can be done only if the comparisons are consistently unfavorable to one twin and show him or her at a marked disadvantage in relation to the other. It is mainly for this reason that the strongest efforts should be made to keep identical twins in balance—physically, mentally, emotionally, socially, and in their achievements—to a greater extent than is ordinarily necessary or possible with two fraternal twins, or with any two singletons in a family.

INTRODUCTION

Sackett's paper attacks the nature-nurture question in a new way. Sackett reasons as follows: If most of knowledge is acquired as a result of our experience in the environment, then a monkey who is raised without ever having the opportunity to see another monkey or to look into a mirror should have no real idea of what a monkey's face looks like. On the other hand, if this important knowledge of how to recognize the face of another monkey were somehow genetically coded, then the monkey might "know what another monkey looks like" even though he had never seen one. He might have been *born* knowing how to recognize fellow members of his species. Sackett's experiment gives a clear answer to this question.

MONKEYS REARED IN ISOLATION WITH PICTURES AS VISUAL INPUT: EVIDENCE FOR AN INNATE RELEASING MECHANISM

J. B. Sackett

Abstract. Monkeys reared in isolation from birth to 9 months received varied visual input solely from colored slides of monkeys in various activities and from nonmonkey pictures. Exploration, play, vocalization, and disturbance occurred most frequently with pictures of monkeys threatening and pictures of infants. From 2.5 to 4 months threat pictures yielded a high frequency of disturbance. Pic-Lever-touching to turn threat pictures on was very low during this period. Pictures of infants and of threat thus appear to have prepotent general activating properties, while pictures of threat appear to release a developmentally determined, inborn fear response.

Research on a wide variety of animals has shown that early experiences can be important determinants of later social and nonsocial behavior (1). Some of these experiences apparently must occur during a limited developmental period if the animal is to exhibit behavior patterns normal for its species. One important, but relatively neglected, area of study in primate behavior involves determination of the developmental importance of different types of sensory input early in life.

The experiment reported here is part of a study examining the effects of visual social and nonsocial stimulation presented to monkeys otherwise reared in total social isolation. The present study asks if totally naive infant monkeys will show differential behaviors toward specific types of visual stimulation, and whether such differential behaviors mature at specific periods during the monkey's development.

Four male and four female rhesus monkeys (*Macaca mulatta*) were reared in individual wire cages (61 by 71 by 71 cm) from birth to 9 months. Three walls, the ceiling, and the area below the wire floor of each cage were covered by Masonite or aluminum panels blocking all visual access to the world outside of the cage. The rear wall of

Reprinted with permission from "Monkeys Reared in Isolation with Pictures as Visual Input: Evidence for an Innate Releasing Mechanism," J. B. Sackett, Science, Vol. 154, 1470-73, December 16, 1966. Copyright 1966 by the American Association for the Advancement of Science.

each cage was a rear projection screen, which also blocked visual access to the outside. The screen had a ground surface, preventing a mirror effect on the inner surface. Thus, reflections of the monkey's own image on the screen were minimized, although slight shadows did appear when the screen was brightly illuminated by the projector. A nonmovable brass lever, 0.6 cm in diameter, projected 7.6 cm into the cage at the bottom right corner of the screen. During rearing, the subjects never saw another monkey, and saw no humans after a hand-feeding period during the first 5 to 9 days of life. Sounds were not controlled, so the monkeys did hear other animals and humans.

On day 14, two types of visual stimulus presentations were begun, as follows.

Experimenter-Controlled Slides

In this procedure pictures of monkeys engaged in different activities and nonmonkey control pictures were projected on the screen. Each individual picture was available for 2 minutes. Six monkey and two control pictures, randomly selected from a large pool, were projected in a daily test session. In these tests the subject had no direct control over the onset or duration of the picture.

Animal-Controlled Slides

In this procedure the subject could expose itself to pictures by touching the brass lever, which operated an electronic contact relay circuit. Lever-touching opened a shutter located in front of a Kodak Carousel slide projector, and also operated a printing counter that recorded the touching behavior. The standard schedule of six

randomly selected monkey and two control pictures was used in each test session. Each picture was potentially available for 5 minutes and 15 seconds. During the first 15 seconds the shutter automatically opened, exposing the subject to the picture that would be available for the next 5 minutes. During the next 5 minutes each lever touch turned the picture on for 15 seconds. If the subject continued to touch the lever, the picture still went off at the end of 15 seconds, and did not come on again until contact was broken and then the lever was touched again.

Up to the fourth month of life all subjects received a minimum of two experimenter- and two animal-controlled tests each week. At 4 months, four subjects received motion picture stimuli, but these monkeys were still exposed to at least one trial of experimenter-controlled slides each week. The motion picture data parallel the data for slides, and will not be treated here.

The 2- by 2-inch (5- by 5-cm) colored slides used as stimuli were grouped into ten categories (four of which are illustrated in Fig. 1). They include pictures of (i) threatening, (ii) playing, (iii) fearful, (iv) withdrawing, (v) exploring, and (vi) sexing monkeys, as well as pictures of (vii) infants, (viii) mother and infant together, and (ix) monkeys doing "nothing." Labels were applied, and stimuli were included for study, only when the pictures received a unanimous title by a panel of eight experienced monkey testers. The final category, (x) control pictures, included a living room, a red sunset, an outdoor scene with trees, a pretty adult female human, and various geometric patterns. Each of the ten categories had at least four different

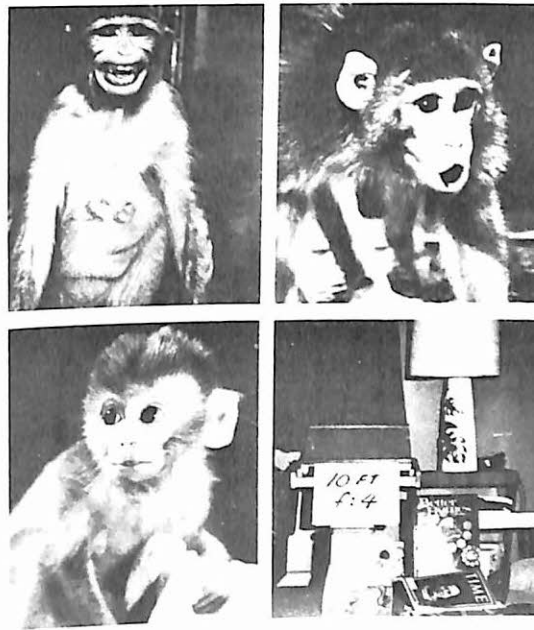


Figure 1

Examples of picture stimuli from the ten categories. The four categories selected for illustration here are (top left) fear, (top right) threat, (bottom left) infant, and (bottom right) control. The actual pictures were in natural color.

examples. Only one randomly chosen slide from each category was used in any individual test session. Complete randomness was restricted such that no single slide appeared more than once in a 7-day period. The monkey pictures were projected with an approximately life-size image.

In animal-controlled tests, the basic data were the number of lever touches producing shutter openings in each of 20 possible shutter openings in each 5-minute stimulus period. In experimenter-controlled tests data were collected on a checklist by an observer looking through a nonreflecting one-way viewing screen in one wall of the cage. These checklist data were frequencies per 2-minute stimulus period of (i) vocalization, (ii) disturbance behaviors including rocking, huddling, self-clasping, fear, and withdrawal, (iii) playing with the picture, (iv) visual and manual exploration of the

picture, and (v) climbing on the walls of the cage. Sex, threat, and aggressive behaviors were also scored, but sexual responses never occurred, and only three threatening or aggressive responses were observed during the 9 months of testing.

Evidence for behavioral differentiation of the ten types of stimuli during experimenter-controlled tests is shown in Table 1. These results can be summarized as follows. First, threat and infant pictures produced the greatest frequency of response on all measures of behavior. All differences between these two types of pictures and the other pictures taken individually were significant beyond the .01 level. Second, there was more vocalization and disturbance with threat than with infant pictures (both $p < .02$). Third, there were no significant differences between the remaining seven monkey pictures (fear, withdrawal, explore,

TABLE 1.
Overall responsiveness to pictures. Mean frequency of the five behavioral measures of all 2-minute periods, during the 9 months of testing.

STIMULUS PICTURE	BEHAVIORS				
	VOCALIZATION	DISTURBANCE	PLAY	EXPLORATION	ACTIVITY (CLIMBING)
Threat	0.61	0.83	0.46	5.0	1.7
Infant	.48	.37	.54	4.8	1.8
Withdraw	.27	.20	.17	2.4	0.8
Fear	.27	.15	.26	2.3	.9
Play	.27	.21	.21	2.7	.8
Explore	.22	.19	.19	2.2	.8
Sex	.17	.24	.18	2.2	.8
Mother-infant	.29	.25	.19	2.5	.7
Nothing	.22	.21	.31	2.7	.9
Control	.14	.12	.16	2.6	.7

sex, mother-infant, and nothing) on any measure (all $p > .10$). Tests of overall differences between any of these seven monkey pictures for each month during rearing also failed to reveal significant variation (all $p > .05$), indicating that these seven pictures were responded to similarly throughout rearing. Fourth, the control pictures had significantly lower frequencies than these remaining seven monkey pictures pooled together on the vocalization, disturbance, and play measures (all $p < .04$), but no significant differences appeared on exploration and climbing activity (both $p > .10$).

The development of responding during experimenter-controlled tests is summarized in Fig. 2. In these data the seven monkey pictures that did not differ significantly from one another in overall frequency of eliciting responses are pooled together to form an "other-monkey" category. Disturbance behavior occurred at a uniformly low level throughout the 9-month period for all pictures except threat. Beginning at 2 to 2.5 months, and peaking at 2.5 to 3 months, disturbance

behavior consisting primarily of fear, withdrawal, rocking, and huddling occurred at high levels whenever pictures of monkeys threatening appeared on the screen. At 3.5 months this apparently innate fear response to threat stimuli declined. The vocalization measure, a response that reflects disturbance but may also reflect contentment, showed a course of development similar to disturbance, except that vocalizations were relatively high with infant pictures. After the first month of life, pictures of threat and of infants received more exploration and play than did other-monkey or control pictures. Interestingly, the first stimulus pictures to receive a relatively high degree of play were threat, which were played with even during periods when disturbance responses to threat pictures were high. Climbing responses (Fig. 3), taken as an index of general activity, exhibited a large increase after the first month. Climbing was most frequent in response to threat and infant pictures from this time until the end of testing. Also illustrated in Figs. 2 and 3 is the finding that all measures for all stimuli generally leveled off at

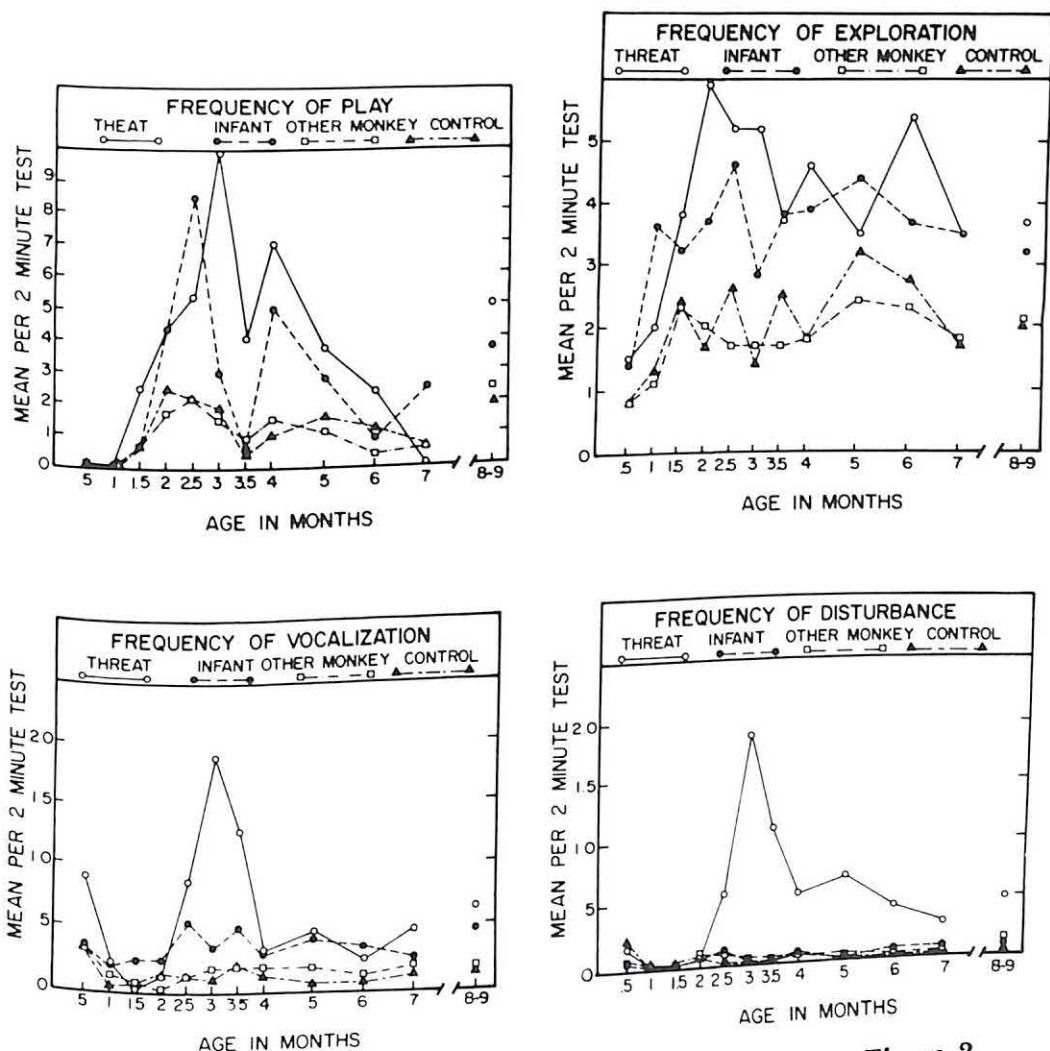


Figure 2

Reactions to experimenter-controlled slides: the development of play, exploration, vocalization, and disturbance behaviors in response to pictures of threatening monkeys, infants, all other monkey pictures pooled, and control pictures.

about 6 months and remained constant thereafter.

Figure 4 presents the data for animal-controlled pictures for all stimuli through the first 6 months of life. Lever-touching was equal for all pictures during month 1. During months 2 and 3 the subjects began exposing themselves more to pictures containing monkeys than to control pictures. About month 3, when the subjects were beginning to show disturbance

to threat pictures in experimenter-controlled tests, lever-touching for threat declined markedly. However, responding for infant and other-monkey pictures continued to increase. During month 4 responding to threat pictures was depressed below the control-picture level, but lever-touching for threat did increase again during months 5 and 6. Lever-touching was generally higher for pictures of infants than for other-monkey pictures from

INTRODUCTION

In the final article in this section Neal E. Miller looks at a wide variety of evidence from his own laboratories and others in summing up the whole range of genetically determined and culturally determined aspects of behavior. Miller speaks with considerable authority, and his overall position is one commonly held today. Both kinds of influences undoubtedly are vitally important. The task for science at this point is to find out what kinds are important where, and whether the limits which seem to be imposed are absolute or merely relative. The answer is not simple, but it does suggest that both sides were at least partially right all along.

PHYSIOLOGICAL AND CULTURAL DETERMINANTS OF BEHAVIOR

Neal E. Miller

The preceding lectures have presented an exhilarating panorama of man's scientific discoveries ranging from the creation of the elements during cataclysmically violent explosions in the ancient vast galaxies of space to the more recent evolution of living sys-

Reprinted with permission, from The Scientific Endeavor by Neal E. Miller, Rockefeller University Press, 1965, pp. 251-65.

tems on our own tiny temperate earth. I have been given the task of making a transition from these topics to those for the next session. How is it that the organisms developing from certain complex helical molecules of DNA will be standing on this platform to talk to you about problems involved in the communication of scientific knowledge, its relationship to public

policy, and the satisfaction of human aspirations?

I cannot answer this question, but I can sketch for you certain relevant steps toward an answer. These will also serve to illustrate a few of the many diverse recent advances in the broad area of behavioral science.

Behavioral scientists are beginning to learn something of how the brain works. The human brain has at least 10^{10} neurons, many of which have thousands of complex branches, resulting in a total of at least 10^{13} (a million times 10 million) synaptic connections. Many of these myriad connections appear to be innately determined. They provide an essential basis for aspects of everyday behavior which all of us casually take for granted. One of these is recognizing a horizontal line. This is not as simple as one might think, because as one's head and eyes move, the same horizontal line stimulates quite different arrays of points on the retina. We are beginning to learn something of how such perception is achieved.

Using microelectrodes to record the activity of individual neurons, Hubel and Wiesel³⁴ find in the visual cortex single nerve cells that fire whenever a specific horizontal line on the retina is illuminated. Presumably the many different receptors stimulated by the points on this specific line are each connected to this cell. Other cells fire when specific horizontal lines in other locations are illuminated. These cells do not respond to general changes in illumination or to small patches on the line. Yet other cells will fire whenever *any* horizontal line is illuminated. Presumably each of these neurons receives connections from many cells each responding to different horizontal lines. Because of the uniformity with which such cells are

arranged in orderly columns, their connections presumably are largely innately pre-determined, although experience is necessary for certain other aspects of complex visual perception.²⁴

Similar arrangements are found for responding to lines oriented in other directions and to a considerable number of other salient features of the visual world. This work begins to give us an idea of how a multitude of specific connections in the brain abstract information from the stimulation of individual rods and cones and allow us to respond to more general attributes, such as a line running in a specific direction.

You have seen how one aspect of the innate organization of the brain provides a basis for everyday behavior. There are many other such examples. Even at the level of human social behavior, we may have many organized patterns, analogous to those which have been studied by the ethologists, which have not yet been recognized as instinctual because they are overlaid by social learning. I agree with previous speakers that innate behavior should continue to be vigorously investigated.

Our understanding of the mechanism of innate organization of the nervous system is being increased by recent research on lower animals. In one type of work Sperry and his associates cut the optic nerve, manipulate it in various unusual ways, and then observe the process of regeneration.^{4, 5} When the eye of a frog is turned upside down, the fibers regenerating from the nerve cell in the retina do not grow along the nearest old path or remain in the same spatial relationship to each other. They are not guided by adaptive function. Instead, each of the 400,000 nerve fibers curves around to establish essentially the

same connection that it had before, producing an orderly visual field which in this case is maladaptively inverted so that for the rest of its life the frog's tongue will snap downward when a fly buzzes by above. Recent experiments on the regeneration of cut optic nerves of fishes have shown that, even when the fibers are forced to detour and pass through abnormal routes in the brain, they will curve back to arrive at their preordained destination.

Such results make a number of older theories untenable and show that various parts of the embryo must differentiate into some fantastically specific coding system with appropriate matches between the more than a million receptors in the peripheral sense organs and the corresponding cells in the brain so that each of the connecting fibers will be guided to grow to the correct destination. Reflection on how a single fertilized egg cell can develop to produce these results gives one a deeper appreciation of the wonderful process of differentiation and orderly growth which has been discussed in Professor Sonneborn's illuminating lecture.⁶⁶

You have had a glimpse of one kind of research that is helping us to understand how the brain is organized and functions in the processing of information. Research on this and other functions is giving us a new picture of the brain. We know that the all-or-none conduction of the long nerve fibers called axones does not apply to the synapses which connect one nerve cell with another and are especially prevalent in the brain. Here the transmitting tip of an axone acts like a gland to secrete a chemical transmitting substance, such as acetylcholine, to achieve a finely graded response which may be either excitatory

or inhibitory. These conclusions are founded on converging evidence from many techniques. Some of this work has been done by biophysicists who thrust micropipettes with several barrels into a single nerve cell, using a conductive solution in one pipette to record the electrical activity of the cell, while minute quantities of various chemicals are injected electrophoretically via the other barrels. Studies with the electron microscope have verified other details. Yet other studies have used a push-pull cannula to wash out and measure for a group of nerve cells the greater production of the transmitter, acetylcholine, when they are active than when they are not.^{18, 22}

The behavioral effects of stimulating considerable populations of cells in certain locations of the brain have been studied by Grossman in my laboratory.²⁵ He implanted a tiny cannula under anesthesia through which chemicals can later be introduced to specific sites in the brain of the normally behaving, unanesthetized animal. His studies have shown that after a rat has been thoroughly satiated on both food and water, injecting a minute amount of acetylcholine or of carbachol directly into a certain part of the brain will cause it to drink, while epinephrine or norepinephrine injected into the same site will cause the same satiated rat to eat. A series of control studies support the most obvious interpretation of these results, namely, that the neuromechanisms involved in the motivations of hunger and thirst are chemically coded.^{26, 52}

Other fundamental changes in our ideas of the brain come from a study of the reticular formation, the central core of the brain which is characterized by a multitude of short interconnections in contrast with the longer

fibers involved in the classical, more peripheral, sensory and motor systems which have long been understood. The study of this system and related nuclei has changed our concept of sleep. We now know that there is a wakefulness region of the brain. If this is damaged, the animal remains permanently somnolent. But there are also two different sleep centers. Instead of activating the animal, stimulation of these centers puts him to sleep. One center is responsible for a distinctive form of light sleep and the other for deep sleep. We no longer think of sleep as an over-all decrease in the activity of the brain. We know that the rate of spontaneous firing of many cells is increased during sleep. The decreased responsiveness to external stimulation seems to be a change in the signal-to-noise ratio, which is produced more by increased noise than by decreased signal.⁴³

The brain used to be thought of as a passive switchboard, activated only when stimulated by the peripheral sense organs. Now we know that each signal impinges on a background of ongoing activity, is modified by this activity, and in turn modifies it. In addition to the pathways coming in from the sense organs are fibers carrying impulses from the brain to the sense organs. These impulses from the brain can either increase the rate of firing of a specific class of sense organs or decrease it. Impulses from the brain can also affect various relay points between the sense organs and the highest levels of the brain. In this way the brain can control its own input.⁴¹

The mechanism I have just described gives us a physiological basis for some of the psychological phenomena of attention, an area in which there has been a great deal of

recent research at the purely behavioral level. This purely behavioral research has discovered many lawful relationships for which the brain mechanisms have not yet been identified, although carefully controlled experiments have demonstrated that many of the phenomena, such as your capacity to listen selectively to either one of two equally loud conversations at a cocktail party, must occur centrally in the brain rather than peripherally in the sense organs.⁹

In addition to its obvious relationship to the eyes, ears, nose, and taste buds that sense the external world, the brain is now known to contain within itself specialized receptors for sensing the internal state of the body. For example, by using a thermode to heat or cool a tiny specific region of the anterior hypothalamus, which is a primitive part of the brain, and by recording from there with microelectrodes, Nakayama, Hammel, Hardy, and Eisenman⁵⁷ have found that the majority of neurons are relatively unaffected by moderate changes in their temperature. However, there are some neurons here that increase their rate of firing when they are slightly heated, and others that increase their rate when they are slightly cooled. These cells seem to serve as, or be connected to, specialized "sense organs" for measuring small changes in the temperature of the surrounding blood.

Heating this region of an animal's brain causes panting and increased blood supply to the skin which serve to lower the body temperature. Cooling it causes the opposite effect of shivering and decreased blood supply to the skin.³⁵ It also stimulates the secretion of the thyroid, which in turn speeds up the body's burning of fuel.²

In the experiments in which only this tiny region of the brain is cooled, these effects produce a fever, but when the whole body is cooled under normal conditions, they serve to restore the animal's temperature to normal. This is one example of a homeostatic mechanism that causes the internal environment to be held at an optimal level, in this case a constant level of temperature.

The regulative effects that I have just described are relatively direct. But the beautiful picture is extended still further by the fact that cooling this region of the brain will make a satiated animal hungry, so that it will eat, whereas heating this region elicits drinking.³ Thus, this temperature-regulating mechanism is tied in with hunger and thirst which motivate behavior that helps the animal to anticipate its needs for the fuel it will burn to keep warm, or the water it will evaporate to cool off.

In short, a whole series of homeostatic mechanisms ranging from changes in metabolism to the motivation of the behavior of seeking food or water is touched off by the cells in the brain that respond to temperature.

A considerable number of other homeostatic mechanisms are known to be involved in the motivation of specific types of behavior. For example, certain receptors in the brain respond to osmotic pressure so that a minute injection into the proper place in the brain of a solution that is slightly more salty than body fluid will motivate animals that have just been satiated on water to drink and also to perform responses that they have learned to get water. Under normal circumstances such drinking reduces the salinity of the body fluids down to its proper level. Conversely, a minute injection of water will cause

a dehydrated animal to stop drinking or working for water.^{46, 58}

Yet other cells of the brain respond to specific hormones so that activities such as nest-building in rats can be elicited by injecting a minute quantity of the proper hormone into the correct site in the brain.²⁰ Hormones can also exert an inhibitory effect on specific areas of the brain without affecting other areas. Thus the progestational steroids normally secreted by the ovary during pregnancy inhibit that part of the hypothalamus responsible for stimulating the pituitary to release the gonadotropic hormone which leads to ovulation. This hormone does not inhibit, however, other parts of the brain which are involved in sexual motivation, sexual performance, or sexual pleasure. A similar synthetic compound is used in oral contraception, a technique which would be considerably less attractive if this hormone's action on different parts of the brain were less specific.⁶³

Motivation has been mentioned a number of times. One of the exciting developments in recent years has been the rapid advance in our understanding of the brain mechanisms involved in certain basic motivations such as hunger and thirst. A combination of physiological and psychological techniques has been especially fruitful. In Figure 1 you will see a rat that has learned when hungry the habit of pushing back a little panel to get food. Since this rat has just been thoroughly satiated on food, it is not performing this habit. But as soon as the experimenter presses a key to stimulate a feeding center in the rat's brain with approximately 15-millionths of an ampere of current, hunger is elicited so that the rat promptly performs the learned habit of going directly to the right place

**Figure 1**

Electrical stimulation in the "feeding area" of the lateral hypothalamus causes a rat, which was satiated and hence previously inactive, to go promptly to the right place and perform the learned habit of pushing back a hinged panel to get the pellets of food hidden behind it. (Photographed in author's laboratory.)

and pushing open the panel to get the pellets of food hidden behind it. He does this repeatedly. As soon as the brain stimulation is turned off, the rat stops performing this habit. Stimulation of this part of the brain can also motivate a rat which has just been satiated on food to learn where to find food. Different tests of this kind prove that such stimulation elicits more than reflex eating; it functions just like hunger in motivating different types of learned food-seeking behavior.⁴⁸

Another advance has been the development of purely behavioral techniques for the laboratory study of non-homeostatic motivations, such as curiosity, the affection of the infant monkey for its mother, and the desire of a human child for social approval.^{6, 23, 29} Progress in such research may be expected to continue; eventually a great deal more will be known about how to stimulate, instead of stultify, intellectual curiosity in the classroom and how to arrange the social situation so that the strong

influence of the classmates will favor, instead of conflict with, the goals of the parents and teachers.^{10, 15}

I have sampled our scientific understanding of how the brain processes information, how it exerts a control over its own input, how it contains sense organs to help it to regulate the internal environment, and how it is involved in motivation. Let me turn now to certain recent knowledge about how we react to stress. Under stress, certain parts of the hypothalamus function as a gland to secrete substances which are carried by the blood stream and excite specific sites on the pituitary gland, causing it to secrete ACTH, which in turn causes the cortex of the adrenals to excrete ACH, which is one of the factors that causes the stomach to secrete hydrochloric acid. Under extreme circumstances, this acid and other effects of the complex reaction to stress are involved in producing ulcers in the stomach. Experimental studies have shown that direct electrical stimulation of appropriate areas of the hypothalamus can produce such ulcers.²¹

Clinical observations suggest that chronic psychological stress can produce ulcers. These observations are confirmed by a variety of experiments in which various animals have been subjected to situations involving difficult decisions, or conflict between approaching a place to get food when hungry and staying away from it to avoid electric shocks. These experiments have shown that such psychological conditions, which presumably produce stress by activating the hypothalamus via the cortex, can elicit the pituitary-adrenal mechanism just described, increase the secretion of hydrochloric acid, and also produce ulcers.^{49, 59}

The same general stress mechanism

seems to be involved also, along with many other factors such as the level of cholesterol and sodium chloride, in the development of arteriosclerosis, and in the production of high blood pressure. Such conditions have been produced by prolonged electrical stimulation of the hypothalamus and also by subjecting animals to psychological stresses which presumably initiate a similar chain of physiological reactions.^{27, 53} Other studies are indicating that cultural conditions producing psychological stress can contribute to the physical symptoms of chronic high blood pressure. Scotch⁶⁴ has shown that members of the Zulu tribe, living in villages in which their culture is relatively unaltered, have a very low rate of hypertension. Those who have been moved to cities where their old culture patterns have been disrupted have an excessively high rate. For both groups, the occurrence of hypertension is correlated with sociological variables presumably involved in stress, but these are different under the two sets of cultural conditions. Somewhat similar correlations between sociological conditions producing stress and hypertension have been suggested by studies of the Negroes in the Chicago area.⁶⁵ These are pioneering studies; their results must be accepted with some caution. More penetrating studies of the roles of sociological variables in psychosomatic symptoms are in progress.

As more work is being done on various aspects of stress, the convergence among the results of experimental studies by physiologists, clinical observations by psychiatrists, experimental studies by psychologists, and cultural studies by sociologists and anthropologists is becoming increasingly impressive. The results from different disciplines are beginning to

fit together into a significant pattern. Additional parts of the puzzle lie within our reach, but we have just begun to grasp them.

Thus far, we have primarily sampled recent knowledge about innate physiological determinants of behavior. As Professors Hutchinson and Mayr have told you, one of the important innate capacities of higher animals is the ability to learn from experience.^{36, 45} Recent advances in our knowledge of learning, memory, and reasoning merit an entire separate lecture, but there is not time.^{11, 14, 16, 19, 32, 37, 42, 47, 50, 51}

The great development of the human brain, which is disproportionately large in the areas representing the vocal apparatus, has been associated with a unique ability to learn and use language. Through language, the experience of one generation is passed on as a cultural heritage to the next. For long-lived organisms, the process of natural selection is extraordinarily slow. Within the individual's lifetime, the process of learning directly from the environment is much faster, but still limited. When the cumulative effects of the lifetime learning of multitudes of different individuals in each generation can be passed on as part of the culture, however, developments proceed with an entirely different order of magnitude of ever accelerating speed.

Man's mode of adjustment to the environment is preeminently cultural. When a lower animal moves into a colder ecological niche, it slowly evolves the capacity to grow a thicker fur; when man moves into a much colder climate, he invents warmer clothes and a better house.

For the ancient men who had a much poorer cultural heritage, the process of innovation was vastly slower. Man has been in existence for

approximately a million years. It is unlikely that his brain has evolved significantly in the last 50,000 years, and perhaps not for much longer. But man did not develop agriculture until 10,000 years ago. It took him about 5,000 more years to invent the wheel as a means of transportation, which as far as we know was invented only once and was never discovered by the Indians of North and South America, in spite of the fact that the Incas had paved roads!¹³ The steam engine as a source of industrial power was invented two hundred years ago, the aeroplane within the lifetime of many of us in this room, and atomic energy and space vehicles still more recently. Compare the glacially slow initial progress with the torrent which you have seen during your own life. This is the effect of the cumulative heritage of culture.

Science is a recent product of this cultural development, one which is powerfully accelerating the process. As President Seitz pointed out during the second session, science flourishes only under certain cultural conditions. Remember what happened to Galileo for advocating the heliocentric hypothesis, how the early anatomists were prevented from dissecting cadavers, and the present activities of antivivisectionists. Modern science would be impossible without an advanced educational system and a strong economic base. Without an enormous cultural heritage, members of the human species would never be giving the papers that are included in the next session on *The Scientific Endeavor*.

To understand man we must study his culture as well as his physiology and psychology. The basic *principles* of human learning and behavior depend on the innate physiological structure of man. But in many cases

we have not yet been able to relate these principles to the innate physiology and have found it most useful to study them at the purely psychological, or in other words, behavioral level, and the discovery of the links to physiology will not eliminate psychology any more than the quantum mechanics of valences has eliminated the discipline of chemistry. On the other hand, the crucial *conditions* of human learning and behavior are being scientifically studied by anthropologists, sociologists, and other social scientists. In order to understand human behavior, one must know both the principles and the conditions.¹⁷ That is why the scientific study of culture is important.

It is often said that human behavior is unpredictable. This is not true. Under appropriate cultural conditions there is a high degree of predictability. Without such predictability, civilization would be impossible. Look around you. It is safe to predict that no one will be sitting there naked. You might stop for a moment to think how much someone would have to pay you in order to make you undress at this meeting. This will give you some idea of the power of culture.

How bad a pest would some elderly relative of yours have to be before you would kill him? That is the strength of the cultural taboo in our society against murder. In India, cows, monkeys, parrots, insects, and other pests ravage the crops of an undernourished population making it extremely difficult to introduce efficient methods of agriculture which involve growing large fields of the same crop and hence invite multiplication of the pests that prey on that crop. Since the Hindu believes that any one of these animals might be the reincarnation of one of his relatives, he will not kill

any of these pests. For him the sacredness of all animal life is more important than his own hunger for food. On the other hand, the Eskimo used to abandon his parents to certain death when they were too old to care for themselves, and the aged parents asked to be killed in this way.⁷ These are but a few examples of the enormous power of the social conditions of culture.

The first emphasis of the social anthropologists was on the extreme diversity of cultures. They proved that an astonishing amount of human nature is the product of culture conditions and hence is not the same everywhere. For example, there are some cultures in which different tribes have interacted for long periods without warfare, and in some cases even without military weapons. Peace is compatible with human nature!^{31, 54, 67, 71}

Even within our own American society, we have various subcultures. In the different social classes, the ideas of what is right and what is wrong, the values placed on education, having an illegitimate child, aggression, saving money, and many other aspects of life are far more different than most of us realize.⁷⁰ There also are regional differences in culture, as illustrated by recent events in the deep South. Certain occupational groupings are minor subcultures. Everyone takes his own immediate culture for granted, and greatly underestimates the degree to which others react differently, because everyone tends to associate with people from a similar cultural background.

As they have advanced beyond their initial descriptions, anthropologists have found that the fact that many radically different ways of doing things in other cultures are quite pos-

sible does not mean that it is a matter of indifference how things are done in a given culture. They have learned much about how the different parts of a given culture dovetail in a lawful, functional way, so that changes of certain aspects may be strongly resisted, and if forced, may produce repercussions throughout the society. To give a simple example, in the Trobriand Islands, instead of a husband supporting his wife, a brother supports his sister, and if his sister has the status of being married to a chief, she must be well provided for. Each chief had many wives which were the source of his economic power, enabling him to supply food and drink for work parties which accomplished essential civil projects such as hollowing out great logs to produce ocean-going canoes. When pressure from colonial administrators caused abandonment of polygamy, the chiefs lost the extra income from having several brothers-in-law, they were too poor to throw work parties, their prestige and authority declined, and necessary civic tasks were neglected. A change in one aspect of the society had unexpected, far-reaching effects.⁴⁴

An additional example of the functional interdependence of various aspects of a culture comes from Murdock's⁵⁵ comparative study of 250 societies which also illustrates the modern trend toward quantitative work. In a statistical analysis he has found that different patterns of sexual taboos in these societies are highly correlated with the ways in which kinship terms, such as "Mother" or "Aunt," are applied to larger or smaller classes of maternal or paternal relatives. The various uses of kinship terms are related to family living arrangements—whether the bride moves in with the husband's relatives, the

husband moves in with the bride's relatives, or both go off independently. These customs are in turn dependent upon economic behavior, for example, hunting *versus* agriculture. Murdock also has shown how those changes which are known to have occurred during the history of these societies progressed in a lawful sequence with the changes in economic behavior coming first, followed by changes in living arrangements, and then in kinship systems.

We have seen that cultural changes are occurring at an accelerated rate. There is less time now for different aspects of a society to adjust by the slow, blind processes of social trial-and-error which were sufficient in the past. Some of the later speakers may be dealing with problems of this kind in discussing new relationships between science, government, and the universities.

In the studies of the hypertension of the Zulus, we saw one symptom of the stress involved in changing from one culture pattern to another. Other effects of the disorganization produced by relatively milder changes are illustrated by the fact that in our own country problems of delinquency and crime have characteristically involved the children of the newest wave of immigrants. Yet other studies seem to show that certain sociological variables, especially social disorganization, are related to a higher rate of mental illness.^{33, 39, 56, 62, 68}

The functional relationships among the different parts of a society become especially important when attempts are made to introduce rapid technological changes in so-called "underdeveloped" countries. The changes occurring in such countries today create an urgent need for us to apply what we know, and at the same time

present unparalleled opportunities for increasing our scientific understanding of the dynamics of rapid social change.

In recent years some political scientists and economists have greatly broadened their perspective by studying other societies and cultures, including some of the technologically underdeveloped ones. Instead of concentrating almost exclusively on official documents or government statistics, they have devised techniques for studying political and economic behavior in the field. Unfortunately, these developments, which are changing these subjects from what they were when most of us studied them in college, are beyond the scope of this paper.^{1, 12, 28, 60, 61}

I shall conclude with one more example illustrating the way in which different disciplines, ranging from the physiological to the cultural, have profitably converged on a behavioral problem. You will remember that culture must be learned by children and that such learning during childhood is a crucial link in passing it on from generation to generation. In addition to this general consideration, clinical observations have suggested that there may be certain critical periods in childhood during which experiences may have an especially profound and long-lasting effect.

These observations have been supported by recent experimental studies on animals. The imprinting, which Professor Mayr⁴⁵ has described, can occur only during a limited period of the newly hatched bird's life. Birds will not learn to follow an object if their first exposure is after the critical period. Female rats restrained from grooming themselves during a certain period of infancy do not show normal nest building in later life and will eat,

instead of care for, their young. Similarly, monkeys reared with artificial mothers do not display normal maternal behavior in later life, and if isolated from other infant monkeys during a certain period of childhood will not acquire normal sexual behavior when adult. They also show other strikingly neurotic symptoms. Bottled sheep do not develop gregariousness, but graze alone. That genetic factors are also involved, however, is shown by the fact that not all mammals become as gregarious as sheep.⁶⁵

The critical periods involve the interaction of developing innate patterns with learning. Their occurrence is firmly established in certain lower animals but we need to know much more about the detailed processes involved, and their occurrence in human infants.

Evidence that there are such critical periods in the development of the human infant has been secured by taking advantage of experiments of nature in which illness or other unusual circumstances have forced the child to be separated from its mother and mother-surrogate. In many, but not all, cases such separation seems to have had serious consequences for adult personality, shifting it toward chronic mistrust, hostility, and delinquency.⁸

Experiments have recently shown that there are critical periods also for psychosomatic effects. It is well known that petting infant rats causes them to grow larger than control litter mates, and also to be superior as adults in certain avoidance-learning tasks. Similar petting at a slightly later age does not produce these effects. At first it was thought that the effects of handling were analogous to those of expressions of parental affection. This turned out to be a far too

flattering misconception of the rat's reaction to being picked up by man. Electric shocks were found to produce similar effects. This and other studies showed that the effects were stressful, involving the links from the hypothalamus to the pituitary to the adrenal that we have already described.¹⁰

Furthermore, biochemical studies of adult animals showed that the brief handling during an early critical period had a permanent effect on the adult animal's physiological reactivity to stress, for example, the production of corticosteroids.

Additional studies are showing that there are critical periods of development during which the injection of certain hormones in minute amounts, which would be ineffective if injected at other times, can have permanent effects upon the rat's physiological development, which in turn have profound effects on behavior.³⁰ Such studies may help us to understand some of the physiological mechanisms involved in critical periods of development.

To return to our previous theme, do the psychosomatic effects of infant stress upon subsequent growth and maturation apply to human as well as to rat infants? One cannot perform experiments on stress with human infants. There are, however, certain societies in which infants at different ages are subjected to certain stressful procedures, such as cutting or burning the skin to form a pattern of scars, piercing the lips or the ears, or molding certain soft bones by pressure. Landauer and Whiting³⁸ have recently made a cross-cultural study of the effects of such procedures on average adult height. They have found that there is a statistically reliable positive correlation between such procedures and mean height. The average

superiority in height of the stressed group is more than two inches, and the data suggest that the first two years are the critical period for children.

Since this is a correlational study, they cannot be certain that there is a causal relationship. Nevertheless, the correlation remains when they analyze the data in such a way as to control for three factors likely to influence height: namely, genetic stock, diet as indicated by protein, and sunshine, which presumably is related to growth-inducing vitamin D.

In short, clinical observations, behavioral experiments on animals, the results of biochemical assays and interventions, and cross-cultural comparisons are converging to show that there are critical periods in infancy during which certain events may have an especially great effect on subsequent physical and behavioral development. The phenomena are real; they are significant; the next step is to discover more about the laws governing them.

In conclusion, current research emphasizes the extreme importance of both innate physiological and acquired cultural factors in human development. We are beginning to learn a few of the fascinating details of the ways in which physiological nature and cultural nurture interact to produce behavior. But the picture is not a simple one, and each new discovery discloses additional tantalizing problems, challenging us to advance scientific knowledge deeper into the vast unknown.

Notes and References

The author wishes to thank Drs. W. Goodenough, G. P. Murdock, C. Plaffman, and H. Riecken for making helpful comments on a rough draft of this paper, and Mrs.

Emily Derow for overtime work in preparing the final copy in time for the meeting.

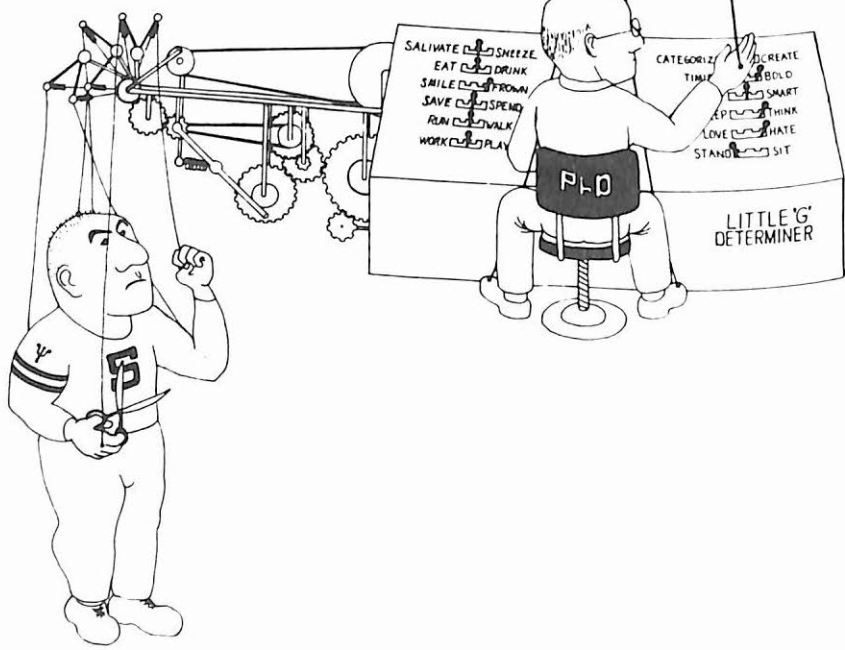
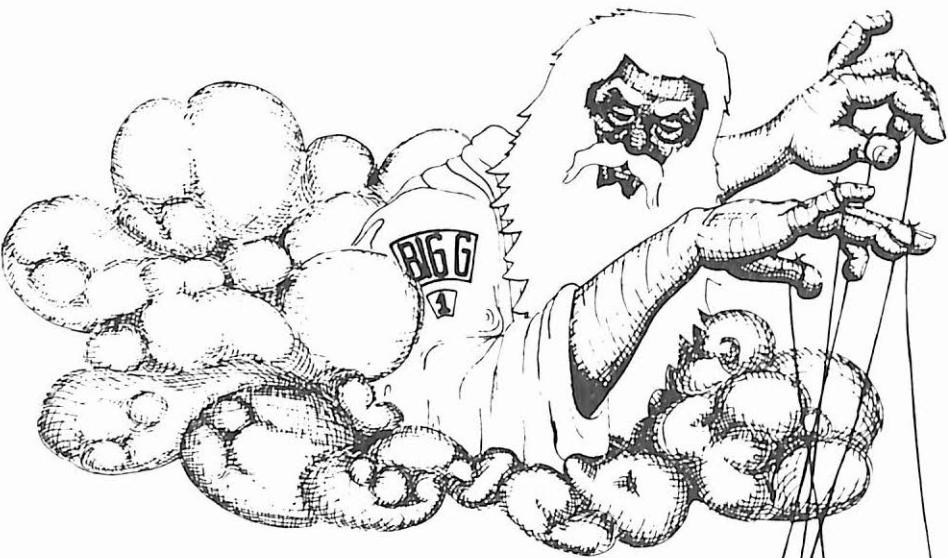
Some of the work on this paper was supported by grant MY 647 from NIMH of the United States Public Health Service.

1. ALMOND, G. A., and VERBA, S., *The Civic Culture*, Princeton University Press (1963).
2. ANDERSSON, B., EKMAN, L., GALE, C. C., and SUNDSTEN, J. W., Activation of the thyroid gland by cooling of the preoptic region in the goat, *Acta Physiol. Scand.*, vol. 54:191-192 (1962).
3. ANDERSSON, B., and LARSSON, B., Influence of local temperature changes in the preoptic area and rostral hypothalamus on the regulation of food and water intake, *Acta Physiol. Scand.*, vol. 52:75-89 (1961).
4. ARORA, H. L., and SPERRY, R. W., Optic nerve regeneration after cross union of medial and lateral optic tracts, *Am. Zool.*, vol. 2:389 (1962).
5. ATTARDI, D. G., and SPERRY, R. W., Preferential selection of central pathways by regenerating optic fibers, *Exp. Neurol.*, vol. 7:46-64 (1963).
- ✓ 6. BERLYNE, D. E., *Conflict Arousal and Curiosity*, New York, McGraw-Hill Publishing Co., Inc. (1960).
7. BOAS, F., *The Central Eskimo*, Bureau of American Ethnology, Annual Report, vol. 6 (1888).
- ✓ 8. BOWLBY, J., *Maternal Care and Mental Health*, Geneva, World Health Organization (1952).
9. BROADBENT, D. E., *Perception and Communication*, New York, London, Pergamon Press (1958).
- ✓ 10. BRONFENBRENNER, U., Soviet methods of character education: some implications for research, *Am. Psychol.*, vol. 17:550-564 (1962).
11. BUSH, R. R., GALANTER, E., and LUCE, R. D., *Handbook of Mathematical Psychology*, New York, John Wiley and Sons, Inc. (1963).
12. CAMPBELL, A., MILLER, W., and STOKES, E., *The American Voter*, New York, John Wiley and Sons, Inc. (1960).
13. CHILD, V. G., *What Happened in History*, Harmondsworth, England, Penguin Books (1948).
14. COFFER, C. N., *Verbal Learning and Verbal Behavior*, New York, McGraw-Hill Publishing Co., Inc. (1961).
15. COLEMAN, J. S., *The Adolescent Society*, Glencoe, The Free Press (1961).
16. DELAFRESNAYE, J. F., *Brain Mechanisms and Learning*, Oxford, Blackwell Scientific Publications (1961).
17. DOLLARD, J., and MILLER, N. E., *Personality and Psychotherapy*, New York, McGraw-Hill Publishing Co., Inc. (1950).
18. ECCLES, J. C., *The Physiology of Synapses*, Heidelberg, Springer Verlag and New York, Academic Press, Inc. (1964).
19. FIELDS, W. S., and ABBOTT, W., *Information Storage and Neural Control*, Springfield, Illinois, Charles C Thomas (1963).
20. FISHER, A., Maternal and sexual behavior induced by intracranial chemical stimulation, *Science*, vol. 124: 228-229 (1956).
21. FRENCH, J. D., PORTER, R. W., CAVANAUGH, E. B., and LONGMIRE, R. L., Experimental observations on "psychosomatic" mechanisms. I. Gastrointestinal disturbances, *Arch. Neurol. and Psychiat.*, vol. 72:267-281 (1954).
22. GADDUM, J. H., Chemical transmission in the central nervous system, *Nature*, vol. 197:741-743 (1963).
23. GEWIRTZ, J. L., and BAER, D. M., *The*

- effect of brief social deprivation on behavior for a social reinforcer, *J. Abnorm. Soc. Psychol.*, vol. 56:49-56 (1958).
24. GIBSON, E. J., Perceptual learning, in Annual Review of Psychology, (A. R. Farnsworth, editor), Palo Alto, California, Annual Reviews, Inc., vol. 14 (1963).
25. GROSSMAN, S. P., Direct adrenergic and cholinergic stimulation of hypothalamic mechanism, *Am. J. Physiol.*, vol. 202:872-882 (1962).
26. GROSSMAN, S. P., Effects of adrenergic and cholinergic blocking agents on hypothalamic mechanisms, *Am. J. Physiol.*, vol. 202:1230-1236 (1962).
27. GUNN, C. G., FRIEDMAN, M., and BYERS, S. O., Effects of chronic hypothalamic stimulation upon cholesterol-induced atherosclerosis in the rabbit, *J. Clin. Inv.*, vol. 39:1963-1972 (1960).
28. HAGEN, E. E., On the Theory of Social Change, Dorsey Press (1962).
29. HARLOW, H. F., and ZIMMERMAN, R. R., Affectional responses in the infant monkey, *Science*, vol. 130:421-432 (1959).
30. HARRIS, G. W., and LEVINE, S., Sexual differentiation of the brain and its experimental control, *J. Physiol.*, vol. 163:42-43 (1962).
31. HENRY, J., Jungle People, New York, J. J. Augustin Press (1941).
32. HOLLAND, J. G., and SKINNER, B. F., The Analysis of Behavior, New York, McGraw-Hill Publishing Co., Inc. (1961).
33. HOLLINGSHEAD, A. B., and REDLICH, F., Social Class and Mental Illness, New York, John Wiley and Sons, Inc. (1958).
34. HUBEL, D. H., and WIESEL, T. N., Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, vol. 160:106-154 (1963).
35. HUMMEL, H. T., HARDY, J. D., and FUSCO, M. M., The thermoregulatory response to hypothalamic cooling in unanesthetized dogs, *Am. J. Physiol.*, vol. 198:481-486 (1960).
36. HUTCHINSON, G. E., "The Influence of the Environment." *The Scientific Endeavor*, New York, Rockefeller University Press (1965).
37. KIMBLE, G. A., Hilgard and Marquis' Conditioning and Learning, New York, Appleton-Century-Crofts, Inc., 2nd edition (1961).
38. LANDAUER, T. K., and WHITING, J. W. M., Infantile Stimulation and Adult Stature of Human Males, *Am. Anthropol.*, vol. 66:1007 (1964).
39. LEIGHTON, A. H., My Name Is Legion, New York, Basic Books, Inc. (1959).
40. LEVINE, S., Psychophysiological effects of infant stimulation, in *Roots of Behavior* (E. L. Bliss, editor), New York, Paul B. Hoeber, Inc., 246-253 (1962).
41. LIVINGSTON, R. B., Central control of receptors and sensory transmission systems, in *Handbook of Physiology* (J. Field, H. W. Magoun, and V. E. Hull, editors), Baltimore, The Williams and Wilkins Co., Vol. 1, section 1, Neurophysiology (1959).
42. LOGAN, F. A., Incentive: How the Conditions of Reinforcement Affect the Performance of Rats, New Haven, Yale University Press (1960).
43. MAGOUN, H. W., The Waking Brain, Springfield, Illinois, Charles C Thomas, 2nd edition (1963).
44. MALINOWSKI, B., Argonauts of the Western Pacific, New York, E. P. Dutton & Co., 63-64 (1922).
45. MAYR, E., "The Evolution of Living Systems." *The Scientific Endeavor*, New York, Rockefeller University Press (1965).

46. MILLER, N. E., Learning and performance motivated by direct stimulation of the brain, in *Electrical Stimulation of the Brain*, University of Texas Press, 387-396 (1961).
47. MILLER, N. E., Some reflections on the law of effect produce a new alternative to drive reduction, in *Nebraska Symposium on Motivation for 1963*, University of Nebraska Press, 65-112 (1963).
48. MILLER, N. E., Some motivational effects of electrical and chemical stimulation of the brain, *EEG Journal*, suppl. 24, Elsevier, 247-259 (1963).
49. MILLER, N. E., Animal experiments on emotionally-induced ulcers, in *Proc. World Congress Psychiatry*, Montreal, Canada, June 4-10, 1961, Toronto, University of Toronto Press, vol. 3:213-219 (1963).
50. MILLER, N. E., Some implications of modern behavior theory for personality and psychotherapy, in *Personality Change* (D. Byrne and P. Worchel, editors), New York, John Wiley and Sons, Inc. (1964).
51. MILLER, N. E., Certain recent developments in experimental psychology, *Proc. Roy. Soc. London, Series B*, vol. 158:481-497 (1963).
52. MILLER, N. E., GOTTESMAN, K. S., and EMERY, N., Dose-response to carbachol and norepinephrine in rat's hypothalamus, *Am. J. Physiol.*, vol. 206:1384-1388 (1964).
53. MILLER, N. E., PEAFFMAN, C., and SCHLOSBERG, H., Aspects of psychology and psychophysiology in the U.S.S.R., in *Some Views on Soviet Psychology* (R. A. Bauer, editor), New York, Timely Press, Inc., 226-229 (1962).
54. MURDOCK, G. P., *Our Primitive Contemporaries*, New York, The Macmillan Co., 110 (1934).
55. MURDOCK, G. P., *Social Structure*, New York, The Macmillan Co. (1949).
56. MYERS, J. K., and ROBERTS, B. H., *Family and Class Dynamics in Mental Illness*, New York, John Wiley and Sons, Inc. (1959).
57. NAKAYAMA, T., HAMMEL, H. T., HARDY, J. D., and EISENMAN, J. S., Thermal stimulation of electrical activity of single units of the preoptic region, *Am. J. Physiol.*, vol. 204:1122-1126 (1963).
58. NOVIN, D., The relation between electrical conductivity of brain tissue and thirst in the rat, *J. Comp. Physiol. Psychol.*, vol. 55:145-154 (1962).
59. POLISH, E., BRADY, J. V., MASON, J. W., THACH, J. S., and NIEMACK, W., Gastric contents and the occurrence of duodenal lesions in the rhesus monkey during avoidance behavior, *Gastroenterology*, vol. 43:193-201 (1962).
60. PYE, L., *Politics, Personality and Nation Building: Burma's Search for Identity*, New Haven, Yale University Press (1962).
61. REYNOLDS, L., *Economics*, Homewood, Illinois, Richard D. Irwin, Inc. (1963).
62. SARASON, S. B., and GLADWIN, T., *Psychological Problems in Mental Deficiency: Part II*, New York, Harper & Brothers (1959).
63. SAWYER, C. H., and KAWAKAMI, M., Interaction between the central nervous system and hormones influencing ovulation, in *The Control of Ovulation* (C. A. Villee, editor), New York, Pergamon Press, 79-97 (1961).
64. SCOTCH, N. A., Sociocultural factors in the epidemiology of Zulu hypertension, *Am. J. Pub. Health*, vol. 53:1205-1213 (1963).
65. SCOTT, J. P., Critical periods in be-

- havioral development, *Science*, vol. 138:949-957 (1962).
66. SONNEBORN, T. M., "The Differentiation of Cells." *The Scientific Endeavor*, New York, Rockefeller University Press (1965).
67. SPENCER, W., and GILLEN, F., *The Arunta*, London, The Macmillan Co., vol. 1:27-28 (1927).
68. SROLE, L., *Mental Health in the Metropolis: Midtown Manhattan Study*, New York, McGraw-Hill Publishing Co., Inc. (1962).
69. STAMLER, R., Acculturation and Negro blue-collar workers, in *The Blue Collar World* (A. B. Shostak and W. Gomberg, editors), New Jersey, Prentice-Hall, Inc., 282-298 (1964).
70. WARNER, W. L., MEEKER, M., and ELLIS, K., *Social Class in America: A Manual for the Measurement of Social Status*, Chicago, Science Research Associates (1949).
71. WILSON, P., *The Tsimihety and the Sihanaka, Human Relations Area Files*, New Haven, personal communication.



PART TWO

free will and determinism

The selections in this part, by Kemeny, Sperry, and MacKay, deal with the very gripping question of whether man can be thought to have free will in a world which is generally viewed by scientists as more or less deterministic. Several crucial questions are raised in these articles. Perhaps the most central is: What is freedom? Is the essence of freedom to *feel free*? If not, what is it? To be unpredictable? A person who always made his decisions by flipping a coin would be very unpredictable but is this what we mean by freedom? Or does it simply mean that no one factor, no outside person, no government has complete control over a human? A robot which did nothing but what his master commanded would hardly be considered free. But, if it were a very sophisticated robot not answering directly to any single master, but rather moved adaptively in a world full of forces impinging upon it, and of concern to it, then that robot, although mechanistic, might appear free at least in some senses of the word.

PRODUCTION

This presentation consists of four short sections from John Kemeny's book, *A Philosopher Looks at Science*. Kemeny is a most remarkable scholar. A student of Albert Einstein, he has made impressive contributions in philosophy, mathematics, and computer science. He was recently appointed President of Dartmouth College, where he has taught for many years. After reading these selections, you will be in a good position to appreciate his fine reasoning and writing abilities. The first section discusses what is meant by a law of nature. Kemeny concludes, somewhat surprisingly, that what is meant is a complete description of all past, present, and future states of the world. Next Kemeny discusses the fundamental problem of causality. He concludes that the relationships now expressed as causal laws can always be expressed in other terms, and that to say Event C "causes" Event E is to say simply that whenever C occurs then E is observed to follow. No more, no less than that this temporal arrangement accurately describes what we have seen in the past. Kemeny notes that if we wish to postulate "causal forces" we may do so but that we do not in so doing increase our understanding. Think about this point a little. In your own experience, when you think that one event causes another do you mean anything more than that the first event invariably is followed by the second?

In the third section, Kemeny discusses the question of determinism, making use of ideas developed in the previous two sections. The final excerpt from Kemeny's book deals with the other side of the coin, the question of what we mean by free will in a human.

A PHILOSOPHER LOOKS AT SCIENCE

J. G. Kemeny

"Always speak the truth—think before you speak—and write it down afterwards."

"I'm sure I didn't mean—" Alice was beginning, but the Red Queen interrupted her impatiently.

"That's just what I complain of! You should have meant! What do you suppose is the use of a child without any meaning?"

There are some fundamental questions a scientist must ask himself before he starts his work. How can he justify his attempts as a scientist? What must he assume about the universe in order to enable him to hope for some kind of success? As a scientist he is going to be called upon to predict the future. He has seen only the past and only a small part of that. What right does he have to assume that the past is a good guide for the future? It is these questions that we must try to answer in the present chapter.

The Law of Nature

The usual argument proceeds as follows: A scientist searches for laws of nature. His search is hopeless if there are no such laws. Therefore the basic assumption of all Science is that there are laws of nature. This sounds very convincing indeed, but we have

been forewarned about making sweeping assumptions. When we state something exceedingly general, the danger is that we may be saying nothing at all. My first major point will be precisely this, that the so-called assumption in this very broad form is empty of factual content.

Of course, no one denies that scientists have formed certain theories which have been referred to as laws of nature. We may be tempted to say that since scientists have found such laws, therefore laws exist. But we know that many of these so-called laws have turned out to be only approximately correct or, indeed, incorrect in many cases. Even a theory so well confirmed as Newton's laws turned out, in the long run, not to hold. Therefore the problem is not that of finding theories which hold approximately in limited range, and of which we are fairly sure, but the problem of laws holding without exceptions, exactly, throughout all time. We may suspect that some of the theories now held are indeed of this kind, but a suspicion is not enough.

Reprinted with permission from A Philosopher Looks at Science by J. G. Kemeny, Copyright © 1959 by Litton Educational Publishing, Inc., pp. 36-53, 187-94, 225-29.

The question is: Can we really assume that such universal laws exist?

What is a law of nature? It is a general statement about the universe which is true. We know that these statements or theories are mathematical propositions which have been interpreted correctly. Thus, our question amounts to asking whether there are any general mathematical propositions which can be so interpreted that they are true? It is my contention that, if we formulate our assumption in this broad sense, we can prove without reference to factual information that there are indeed laws of nature. In other words, I will maintain that the so-called assumption is true, but analytically true.

It is a great pity for the Philosophy of Science that the word "law" was ever introduced. The usage of "law" carries the connotation that it can somehow be disobeyed. The idea of obeying or disobeying should not have entered these discussions at all. I am afraid this is part of our tendency to try to recreate the universe in our own image. It may perhaps date back to the primitive idea of God. To primitive people, God is but a very large and powerful human being. We even find examples where they try to make deals with Him. This may be in the form of sacrifices or promising to be good if he will only do them a certain favor. In mythology we find many examples where the gods are endowed with human strength and human weaknesses. It is perhaps a remnant from these days that we associate the idea of a law with nature. Presumably the laws of nature were passed by God. They were laid down on a heavenly piece of paper and these are the things that we must do. These are the things that nature must obey. There are even recorded examples of

violations of the laws of nature, and these are called miracles.

We must realize that nature is not like a human being. Nature cannot obey or disobey. The laws of nature do not prescribe but, rather, they describe what happens. A law of nature is but a description of what actually takes place. In contrast to this, whenever a human law is passed, then it must be possible to violate this law. When we pass a law forbidding murder it is, of course, assumed that it is humanly possible to commit murder. If it were impossible to commit murder, then there would be no point at all in passing a law against it. Therefore we admit the possibility of someone acting against the law and we have to invoke penalties and enforcement agencies to see that the law is obeyed. But a law of nature cannot be disobeyed. There is no need for any kind of enforcement agency. The reason it cannot be disobeyed is much simpler. Since a law of nature is no more and no less than a careful record of what actually happens, there is no possible way of violating it. If in any experience one keeps a careful record of all that actually happens, then in no sense can this record be violated. After all, if we have stated exactly what has happened, then nothing can ever be done to undo this.

As a matter of fact, from this point of view a law of nature is a very dull thing. It is like a court record of proceedings on an average hearing before the judge. The only thing that lends it glamour is the fact that it is a record not only of the past, but also of the present, and most important of all, the future. It is like a court record of all past, present, and future proceedings. Anyone who is in possession of such a record will know what lies ahead of him.

Some years ago Hollywood produced an interesting movie on this subject. It dealt with a man who is given a newspaper each day. But by some unexplained method this paper turns out to be tomorrow's newspaper. Therefore, each day he knows what tomorrow's headlines will be. Naturally this gives him tremendous power over his fellow human beings. He can go to horse races and win any amount of money that he pleases. He can use the information available to him to get a good position and to advance himself and his friends. Of course, there has to come a moment in a Hollywood movie when the hero is punished for all this undeserved wealth. One day he reads that tomorrow's headline will be a report of his own murder. He does everything possible not to get near the place where the murder is supposed to take place, but he does not succeed. Through a strange series of circumstances he ends up in just that building and at just the right time. A shooting takes place and he is accidentally shot though not killed. But he is reported dead, and indeed the report of his death is tomorrow's headline.

The laws of nature are, in a sense, like tomorrow's newspaper, but they report not only what happens the next day, but a week from now, a year from now, or many millions of years from now. In addition, the laws of nature are like the perfect newspaper. They never report anything mistakenly and they always include every last detail in their report. Any man in possession of a law of nature would have the most tremendous power ever dreamed of by mankind. It is for this reason that scientists spend their entire lives trying to get a small part of tomorrow's newspaper.

We are now in a position to prove

the existence of laws of nature. Actually we can prove something stronger than that. We can prove that any given phase of the world is covered by some mathematical law correctly interpreted. Let us select an example that is supposed to be particularly unpredictable. Naturally I mean the weather. To make the example more specific, let us take the temperature of New York City. First of all, we must select a suitable mathematical language. But this has already been done for us. We will state our law in degrees Fahrenheit, the usual measure of temperature. What form would the law have to take? It is supposed to state for any moment in the past, present, or future just what the temperature was at the place that is now New York City. Such a correspondence is known in mathematical terminology as a function. More specifically, we have to describe the degrees Fahrenheit in New York City as a function of time. This is symbolized by writing $f(t) = d$. The question then is: Is there any mathematical function which, interpreted in the way we have just indicated, correctly describes the temperature in New York City at various times? But as soon as we have stated the problem this clearly it is very easy to see that such a function exists. Given any moment of time, there is no difficulty, in principle, in recording the temperature in New York City. Again, in principle, there is no difficulty recording this for all moments, past, present, or future. And this record defines a function. Of course, in practice we could never carry this out, but this is relevant only to the question as to whether we can find out what this function is, not to the question as to whether such a function exists. The possibility of making such a record,

in principle, is sufficient to guarantee that this is a well-defined mathematical function. Hence it proves that this aspect of the world is covered by a law of nature. A similar argument will work for any other phase of the entire universe.

Once we have seen that the existence of a law of nature is no more than the theoretical possibility of keeping a careful record, we can do much more than we have done so far. It is possible, in principle, to keep a record not only of one event, but of any event in the entire history of the universe. If we are then prepared to imagine that some all-powerful heavenly agent keeps a careful record of all events in the universe, then these records, together, would form a law which covers everything that happens in the universe. Let us call this record *The Law of Nature*.

Naturally it is not claimed that there is only one possible way of constructing *The Law of Nature*. Since this law is not a pure mathematical law, but an interpreted one, there are many different ways of connecting mathematics with the universe. For any choice of an interpretation we will have one version of *The Law of Nature*. But since any one version is as good as any other one, we will suppose, for the sake of the argument, that one definite version has been selected and that this is our standard reference book in terms of which we can discuss the basic assumptions of Science.

We have now reached the surprising conclusion that there is no need to assume anything about the existence of the laws of nature. Then why have so many eminent people stressed the necessity of making such an assumption? The only reasonable conclusion is that, although the assump-

tion usually found in books is analytically true, there must be a slightly different version of this assumption which makes a good deal of sense. Indeed we will find that this is the case.

Causal Laws

Let us return to our original problem of trying to find factual content for the empty assumption about the existence of laws. One promising line is to assume not only that some laws exist, but that laws of a special form exist. We will consider six possible forms of laws and hence six new kinds of possible assumptions: (A) Causal laws; (B) time independent laws; (C) conservation laws; (D) minimum principles; (E) continuous laws; (F) laws involving a given number of dimensions. We will show that of the six forms mentioned above, four lead to empty assumptions while the last two lead to genuine factual assumptions.

It will help in these considerations to illustrate our points with a specific example. Let us consider the law governing a stone that is thrown into the air. If we throw it from a height of h feet with a velocity of v feet per second, then its height at future moments is given by the formula $d = h + vt - \frac{1}{2}gt^2$. In this t is time measured in seconds, and g is a fixed number (approximately 32.2). This law, essentially due to Galileo, is of considerable interest since, if we neglect air resistance, the law applies to any physical object, be it a heavy stone or a feather. We will eventually show that this law can be rewritten in any of the four forms A through D.

The search for causal laws is deeply tied up with our subconscious tendency to recreate the universe in our own image. In our lives we search for

motivating forces for all our actions. We say that a man is irrational unless he has had a "good reason" or a "good cause" for what he did. We are constantly stating what caused us to act in the way we did. "I went to the ball game because my father taught me to love the Dodgers." "I voted the Democratic ticket because of the speech I heard in October." "I believe in the U.N. because I believe it is our only hope for peace." In all three cases something happened in the past that caused the action in the present. What typifies causality is the fact that an event taking place at some time somehow brings about an event at a later time.

We demand that nature should follow our example. If anything at all takes place we immediately search for the cause in a previous chain of events. Why does the tree fall? Because we cut it down. Why does a balloon rise? Because air pressure lifted it. Why does wood burn? Because oxygen combined with it. In each case we search for an event which brought about the result. Subconsciously we feel that nature was forced by its past history to take a certain course of action, the same way as we as individuals are forced to take certain steps. When primitive people personify nature we laugh at them. Yet in many ways we are much more laughable, since we ought to know better.

One of the foremost British philosophers, David Hume, deserves credit for first pointing out that our attitude toward causality must be changed. Let us consider a typical example. We pull the trigger of a rifle and there is a loud noise. The cause is the pulling of the trigger and the resulting noise is the effect. If this happened but once in the history of the universe we would never associate the two

events with each other. The reason we say that the former caused the latter is that we have often observed these two events together. Whenever we pull the trigger of a rifle (naturally a rifle in good working condition, loaded, and without a silencer), there follows a loud noise. But this sentence contains everything that is of factual significance. Whenever one event takes place it is always followed by an event of the second kind. If we add anything to it about one causing the other one, there is certainly nothing in our experience that warrants the assumption. The statement that an event of type A is always followed by an event of type B is a true general proposition and hence is a law of significance to Science. For example, if we pull the trigger in the future under the proper circumstances, we can predict that a loud noise will follow. There is no need whatsoever to assume that some mysterious force connects the two events. And even if there were such a force, if we cannot observe it in any way, we do not care about its existence. Hume does not object to calling one event a cause and the other one the effect, as long as it is clearly understood that this means no more and no less than that A is always followed by B.

What typifies causal laws is that, if we are furnished certain information about the present moment, the law will give us information about the future. This, of course, is of tremendous significance for Science. There is a second point of view from which causal laws are very important. They are the key of the division of labor between the theoretical scientist and the experimentalist. It is the role of the theoretical scientist to find causal laws, whereas it is the role of the experimenter to furnish data about the

present. Combining the labor of these two types of scientists we are furnished information, not only about the present but about the future as well.

Let us now return to our example. The law we stated governing the tossed-up stone is in explicit form giving us information about past, present, and future. But it is easy to change this into a causal law. Suppose we state merely that the acceleration is equal to $-g$. (The acceleration is the rate at which the velocity changes. The minus sign indicates that as the stone is flying upward its velocity is decreasing, while as it falls to the ground it falls faster and faster.)

It will be objected that this is not a causal law, since no "cause" has been found. But nothing is easier. We will say that the acceleration is caused by a "gravitational force," that is, by a force that causes the gravitational acceleration. How do we observe the force? By observing the acceleration. How strong is the force? Just strong enough to bring about the acceleration. This may seem artificial, but this is exactly what Newton did. In every type of motion we can observe the acceleration, a , and the mass of the moving particle, m . He then postulated a force causing the acceleration, of strength ma . Since in many types of motion the acceleration is inversely proportional to the mass of the moving particle, this gave Newton a constant force governing the motion. In our example even this satisfaction is lacking. The acceleration is independent of the mass of the falling body! But there is no difficulty in overcoming this; if the acceleration is constant, then we need simply postulate that the gravitational force is proportional to the mass of the falling body. Together with the fact that the force has

strength ma , this says no more and no less than that the acceleration is a constant. The only possible justification for this bit of sleight of hand is that we demand that our laws be in the form of causal laws.

From this new form of the law alone we can say nothing at all about the position of the stone at any given moment. If we carry out a simple mathematical operation, namely integrating twice, we arrive at the explicit form of the law we started with, but with two important gaps. We do not know the value of h and we do not know the value of v . These values must be found by observation. Once these two numbers are found we have our original version reinstated, and hence we can have complete information about the flight of the stone. No matter how many stones we wish to study, for any stone we will need only two numbers. If we want to study a million objects thrown up into the air, the observer will have to get us two million numbers. As far as laws are concerned the same causal law will suffice. We have said that a causal law determines the future given the present. The two million numbers are our description of the present. The causal law does all the rest of the work for us. We might picture the process intuitively as follows: The given numbers tell us where the stones are at this moment and how fast they are moving. The causal law tells us then what will happen a brief time interval later. And given the information about this moment the causal law again tells us what happens after a short time has elapsed, etc. Putting all this information together we get a complete picture of the future.

We now see the great advantage of using this particular form for a law. Instead of having separate laws for

stones thrown from different heights and with different velocities, we need but the single, simple law that the acceleration is equal to $-g$. All the rest is left up to the experimentalist. The more data he collects about the present, the more we will be able to say about the future. But it is also clear that this is but a convenient form for the law; it is not in any sense a special kind of a law. We had no difficulty at all in changing our law into a causal law, and it can be shown that there is no difficulty in this in any conceivable law. Therefore the so-called assumption about the existence of causal laws is not really an assumption but a convention as to the form to be used for laws of nature. We will later see that assumptions of the form B, C, D also are no more than conventions.

Naturally, conventions are very useful in any undertaking. The danger is in mistaking a convention for a factual assumption. There have been many philosophers who have tried to take the fact that scientists use causal laws and draw far-reaching conclusions from them. It is clear that the only legitimate conclusion to be drawn from the existence of causal laws is that scientists like to put their laws into causal form.

It has been maintained that the principal competitor of a causal law is a teleological law. The latter is a law which determines the present, given the future—rather than determining the future. A great deal of time and space in publications has been devoted to a discussion as to which type of law is superior. I do not want to leave this topic without considering the question.

Let us construct a causal law. We will choose the weather as our subject, and simply consider whether it

is "good" or "bad." As the "past" let us consider the last three days. The law states that if of the last three days *exactly two* were good, then the weather remains the same as yesterday's; otherwise we have a change in the weather. We will see that this is a complete causal law as far as the goodness of the weather is concerned.

Suppose that we have had two good days followed by a bad day. We can symbolize that by "GGB."

Last three days.	GGB
Exactly two days were good, so the weather stays the same (i.e., bad).	GGBB
Now we see that of the last three days only one was good—a change is due.	GGBBG
Same as above—another change.	GGBBGB
Same as above—another change.	GGBBGGB
Now two of the last three days have been good—no change.	GGBBGGBG
Same as above—again no change.	GGBBGGBGG
Now all three days have been good, hence we are due for a change.	GGBBGGBGGB

We have now returned to our original case, and from here the weather repeats itself. We are led to a seven-day cycle. The information needed, namely the weather of the last three days, is the initial data, and everything about the future follows from this by means of our miniature law.

Can we say anything about the past? We again start with GGB. Let "X" stand for the weather of the day before, giving the sequence XGGB.

Since XGG was followed by a change in weather, XGG cannot have *exactly* two good days in it. Hence $X = G$. Let "Y" stand for the day before, in the sequence YXGGGB, or YGGGB. YGG is followed by no change; hence exactly two of these three days must have been good. Therefore, $Y = B$. And so on. Given the initial data, our law determines not only the future, but also the past. This is true also of the differential equations that express the causal laws of physics. What does this do to the old philosophical controversy about whether laws should determine the future on the basis of the present, or determine the present on the basis of the far future? The entire controversy seems to have lost its meaning.

Determinism

"If any of them can explain it," said Alice, "I'll give him sixpence. I don't believe there is an atom of meaning in it."

"If there is no meaning in it," said the King, "that saves a world of trouble, you know, as we needn't try to find any."

This quotation expresses my sentiments concerning the problem of determinism. It seems to be the most controversial pseudo-issue ever made important by dressing it up in Big Words.

METAPHYSICAL DETERMINISM

The problem is whether the future is or is not determined. I must confess that I have no idea what this means, and the more I study various papers written on this subject, the less I understand what an undetermined future would be. It is admitted, on the one hand, that the future will take

place in one definite way and, on the other, that it has not yet taken place. Beyond this very little can be added.

"The Dictionary of Philosophy" gives us the following definition of determinism: ". . . that every fact in the universe is guided entirely by law." This is typical of the type of confusion that is necessary to deceive us into thinking that there is a problem here. We must believe that laws "guide" facts, or that laws "determine" facts. If, instead, we recognize that laws only serve as descriptions of facts, then it becomes clear that The Law of Nature describes all facts—hence that the universe is determined in this sense; it is also clear that this is a characteristic of all conceivable universes.

Another usual formulation is to ask whether every event has its cause. But we have seen that whether we describe the laws in causal form or not is entirely our free choice. Causal laws are very useful, because they lead to a division of labor. We try to formulate a law which, given the present, will determine the future for use. Then the more facts we can accumulate about the present, the more we will be able to predict concerning the future. Discussions of determinism are intermixed with discussions of our ability to find such laws, and of being able to find *all* the initial conditions.

Far be it from me to deny the importance of the question as to what degree we humans can know the future. But whether or not we humans can in principle predict the future is a question about our limitations, not about the nature of the universe. To state the fact that we can never predict the future completely in the terms "the universe is not determined" is not only bad syntax, but it leads to atrocious Metaphysics.

Let us conclude that everything relevant to this pseudo-problem has already been stated . . . [earlier] . . . and pass on to a more interesting question.

Predicting the Future

Before getting lost in the problem of the predictability of the future, let us ask who is to make the predictions. Is it to be a super-human intelligence, which has unlimited knowledge of The Law of Nature, or is it to be an ideal human being, or is it to be an actual person?

On the one extreme, a super-human intelligence can "read off" the future from The Law. But this is not surprising, since The Law is a description of past, present, and future. At the other extreme, it is obvious that at the present, with our present laws, we are unable to predict the future with any accuracy. So we see that any interesting question must lie somewhere between these extremes.

We can distinguish four factors influencing the problem: (1) The laws available. (2) The facts available. (3) The reasoning powers available. (4) The time available. As to the first factor, we have already seen that we should allow ourselves less than The Law of Nature. We may suppose that the law available is a consequence of The Law—that is, describing some part of the history of the universe—or that it is an approximation to such a partial description. Instead of taking it to be a complete description of a part of the universe's history, it is more realistic to take it to be a causal law applying to certain kinds of phenomena, which, coupled with facts about the present, predicts these phenomena in the future. The problem for such laws is to know just what

area can be covered by them. One school maintains that such laws will be found to cover all types of phenomena. A more conservative position is that, although this goal will never be reached, we will continually broaden the field to which such laws are applicable.

On the other side we find various views maintaining that some areas will never be describable by human laws; some of the positions will be discussed in the following chapters. Again it can be argued that, although causal laws can be found, they are always only approximately true and that, besides broadening the field of applicability, we must also try to get better and better approximations to the truth. It is even argued frequently especially since the success of Quantum Mechanics, that our more precise laws cannot be causal, but must be statistical. (This last distinction is, for practical purposes, not as fundamental as it may seem, since we have already noted that due to errors in observation we can only make statistical predictions from any kind of law.) Disputes between these various positions are fruitful, and they influence significantly our beliefs as to the predictability of the future.

Let us next turn to facts. Even if two scientists agree as to what we can expect from our theories, they may disagree as to the practicability of predicting the future, on the basis of varying views as to the possibility of finding the necessary facts to apply the theory. Newton's laws will tell us all about the future of the solar system, if we know just where the planets are at a given moment in relation to the sun. Errors in prediction may be due, on the one hand, to the inaccuracy of the law (which became apparent in discrepancies in the orbit

of Mercury) and, on the other hand, to inaccuracies in determining the positions of the planets, which can be reduced but never entirely eliminated. These errors "spread" in time, which is why short-range predictions are more reliable than long-range ones. It is well known to anyone who has to do a lot of calculations that the so-called rounding errors may become very significant in a long computation. Similarly a very small error in calculation may be magnified as the result is used over and over again. Introducing a small error of observation is like introducing a calculating error into the calculation of predicting. The more we use this figure, the greater the error may become. Perhaps a good analogy is provided by the taking of a drug. If the druggist makes up the prescription just slightly too strong, it may have little effect at first, but it may eventually lead to death instead of a cure.

The problem of predicting is further complicated by the celebrated Uncertainty Principle. When we come to atomic and subatomic phenomena, we find that our methods of observing and measuring are so crude that we cannot help disturbing the system to be observed. When we measure the length of a table, we put a yardstick up against it, which act pushes the table. But this push is negligible. We can further reduce this push by not touching the table, but only sighting it. This, of course, involves reflecting some light from the table. This light ray also exerts a push, but so slight that no one would ever take it into account. However, when we come to measure minute particles, we are like elephants trying to measure a violet. According to the Uncertainty Principle there is an absolute limit to the accuracy we can achieve. This shows

that, whether our laws are statistical or causal, there are severe limits on the accuracy of our predictions, especially long-range predictions.

The third factor, that of limitations on our powers of analysis is not nearly as frequently mentioned. Yet there are famous examples to illustrate this difficulty. To derive predictions from a given theory often requires long and extremely difficult mathematical arguments. Given Newton's laws, it is easy to show how two bodies move relative to each other, each attracting the other one. But when we ask the same question for three bodies, the answer has still not been found in complete generality. We have all kinds of approximations for special cases, but no general solution. In the case of the solar system the motion of two planets around the sun is solved by assuming that the attraction of the planets on each other is nearly negligible compared to the pull of the sun. We first calculate the pull of the sun and the resulting planetary paths; then we introduce a correction for the mutual attraction of the planets. This method necessarily introduces new inaccuracies, which are vastly increased when we consider dozens or hundreds of billions of bodies.

The extreme example of this difficulty in calculation is given by Einstein's recent work. As we have noted, his Unified Field Theory, published in 1950, is entirely untestable for the time being, because the mathematical problems that need to be solved are beyond our ability. In this case Science must wait for major progress in Mathematics before a single prediction can be made from the new theory.

The last of the four factors, the time element, does not affect predictability in principle, but it does affect

the practical problem. It is very nice to know that with our present theories we can predict tomorrow's weather. But if it takes our calculating machines a full month to carry out the necessary prediction, then it is of very little practical use. We could always argue that with the progress of machine-construction the prediction will be fast enough to be helpful. But there still remains the possibility that some predictions by their nature require more time than it takes for the event to happen. This may, for example, be the case in trying to predict human decisions.

In summary, let us recall that this problem of predictability has nothing to do with the problem of whether the universe is in some mysterious sense determined. Rather it has to do with our human abilities to comprehend the universe. At any one stage it is a scientific problem to see how far and how accurately the future is predictable by our limited means. But it is up to the philosopher to speculate which of these limitations are essential and which will be removed, or at least improved, with human progress.

Two Related Topics

Let us first consider an ingenious argument to show that complete prediction of the future is in principle impossible. Its source is not entirely clear to me, as I have heard it attributed to various authors. Recently K. R. Popper studied this problem exhaustively.

Suppose that we have found the causal laws describing physical nature, or even that some spirit whispers them in our ears. Then we can build a machine which is specially designed to predict the future, given data concerning the present. Let us build such

a machine, and put it in a laboratory which is carefully isolated from the outside. Let us wire the machine so that it will calculate the future for 100 years, or at least so that it can answer any question about the next 100 years. Suppose the machine has a certain signal indicating that the answer to our question is "Yes" and another for "No." Let us connect to the machine an electric fan, built so that it could run for hundreds of years if not shut off. Let us connect this fan so that the only way it will be shut off is by means of a "yes" signal from the machine.

Let us now suppose that we somehow supply the machine with the needed data concerning the present. It doesn't matter whether we did this by measurement, by accidentally guessing the truth, or even had a superior spirit hand it to us. Somehow the machine is given all the needed data. Then we can ask it any question concerning the next 100 years. Let us ask it: "Will the fan be running 99 years from this moment?" The machine goes into an intricate calculation, and if the future is predictable, it should sooner or later come up with an answer. If this is to be any use at all, it must arrive at an answer in less than 99 years. Suppose it arrives at the answer "Yes," telling us that the fan will be running. Then this very answer causes the fan to stop! But if the machine says "No," then the fan will keep on running for several hundred years. So in either case the machine comes up with the wrong answer.

It may occur to you that the trouble in this argument is that the machine is part of the system studied. But, alas, we too are part of the universe about which we make predictions! If we analyze the situation, we see that there is nothing wrong with the argu-

ment. So at least one of our assumptions must be wrong, and we see that these correspond to the four factors affecting predictions. First, we assumed that we somehow have the laws governing such physical problems. Perhaps these laws cannot be made comprehensible to us, and hence to our machines. Secondly, we assumed that somehow we had all the initial data necessary available to us. Perhaps this is beyond us. Thirdly, we assumed that the machine can carry out the necessary calculation, while it may be the case that such problems prove unsolvable to it. Finally, and perhaps most simply, it may be the case that it always takes the machine more than 99 years to finish the calculation, in which case the contradiction is removed. The machine sends its "yes" answer in the hundredth year. This turns the fan off, but it does not alter the fact that the fan was still running after 99 years. Any one of these factors or all of them could make the required prediction impossible, and it would be most interesting to determine just which is the actual cause.

The second topic relates specifically to the difficulty in solving the mathematical problems involved in predictions. K. Gödel proved a most interesting theorem to the effect that no matter what mathematical methods we have at our disposal, some of the questions we can ask ourselves cannot be answered by these methods. This does not preclude that with the progress of Mathematics a given unsolvable problem will be solved; however, then there will be new unanswerable questions. In other words, the progress of Mathematics is an infinite progress, where at each stage there remain unanswerable questions.

We know that each body of theories in Science is an interpreted mathemat-

ical system of a fairly advanced type. Hence it has some unanswerable questions in it. What we don't know is whether any of these unanswerable questions concern predictions. But we have no reason to suppose that none of them do. It is entirely possible that even if we knew exact causal laws applying to all types of phenomena, and if we had an unlimited supply of data, we would still have to wait for mathematical progress for certain predictions. For some reason this possibility seems to have been generally overlooked. When we find some Physics problem as yet unanswered, we assume that it is because no one has yet found the right "trick" to solve it. Actually it is possible that this problem needs an entirely new branch of Mathematics.

This suggests the interesting possibility that, not only must we expect an endless advance on the part of the theoretical scientists, but that this progress must go hand in hand with the endless progress of the mathematicians, if it is to be of use to mankind.

Free Will

Any treatment of the mind-body problem which ignored the problem of free will would be severely criticized. Are we really free to make decisions, or is this an illusion?

The famous arguments that there is no free will remind me of an old anecdote. A country hick takes his son to the city zoo. His son is full of curiosity and keeps asking his father the names of these strange animals. The father has a difficult time, because he can't read, and so is forced to keep guessing. Just about the time that the parent's imagination runs out his son spots a giraffe. "Dad, what kind of animal is that?" The father thinks, and thinks, and stares in more and more amaze-

ment, until he exclaims: "Son, there ain't no such animal!"

In the last analysis no amount of ingenious, involved, and incomprehensible argumentation will convince any of us that we have no choice as to whether we use our right or left hand to scratch our ear. It is a fact that we do make free decisions daily, and the role of Science is to explain this fact, not to explain it away.

Then what is the problem? One argument is that our will cannot be free, since the universe is determined. If the future is already determined, how can our supposed choice affect it? . . . we decided [earlier that this sort of determinism] is a pseudo-problem which arose out of a misunderstanding of the nature of scientific laws. We could restate the argument by saying that we cannot have a free choice because the Law of Nature says what the outcome of our choice will be. If it is already "written," then we have no real choice. The Law is not something binding, but a simple description of all events, past, present, and future. Among other things it describes how we choose. This is the only reason why our decision must be in accordance with it. It would be just as correct, and perhaps less misleading, to say that the Law of Nature depends on our choice, instead of the reverse.

. . . we saw that this pseudo-problem of determinism is closely related to the very real problem of predictability, and this is the more reasonable version of the free-will problem. It is generally supposed that if a scientist knew enough about a person, he could predict what the person would do under given circumstances. If an outsider can do this, how can we say that we had a real choice? How can we be free to choose, if the scientist will

tell us ahead of time what we will choose?

First of all, it is highly questionable whether human actions are fully predictable. But this problem we will postpone until the following section. Let us suppose for the sake of argument that our actions are predictable. If the scientist making the predictions could simply watch us from the outside and yet tell us what we are going to do, this would be surprising indeed. But this is not what is maintained. What *is* maintained is that if the scientist knows all about our physical make-up and past history, that is if he knows all internal and external factors influencing our decision, he will be able to predict the outcome.

As we have seen, in predictions all relevant data must be known, besides having the necessary theories. So even if the materialists are right in supposing that mental processes can be reduced to nerve actions, and even if these theories were completely known, he would still have to know the initial data—external and internal. But then he would know all the facts that we know and would know the machinery we use for making decisions (remember we are assuming the materialist thesis for the moment, so that knowing our brain means knowing our mind), and then it is not at all surprising that he can predict.

I know my best friend very well, and it is entirely possible that I can guess his decisions with good accuracy. Does this mean that he had no free choice? For this I did not even need to know all the relevant facts, nor know any far-reaching theories; all I need are some good guesses based on past experience. After all, what is a scientific prediction but a good guess based on past experience? Looked at from this point of view the argument reads

"You are not free to make a decision, because I can frequently guess what you are going to decide," which is ridiculous.

In short we do have free will. The only problem is to find out the mechanism by which we make decisions. This is a difficult and important problem for the psychologist, but not for the philosopher.

Are Human Actions Predictable?

Let us turn to this very interesting question. The free-will dispute disturbs the average college student for two reasons. First, he is afraid that someone will prove his freedom an illusion; and secondly, he feels that predicting his actions is an inexcusable disturbance of his intellectual privacy. The former we have seen to be without point, but the latter is a real "danger."

Instead of asking flatly whether our actions are predictable, let us rather consider to what extent they are predictable. Certainly some actions can be predicted. Put a steak in front of a starving man, and you will have no trouble predicting his actions. On the other hand, we certainly cannot predict all actions, at least not for the time being.

There are two fundamentally different ways of trying to form theories about the actions of a given man. One is the behaviorist approach of looking at him from the outside only and forming generalizations about the pattern of his behavior. This is certainly the easier approach, and more fruitful in the short run. The second approach is to try to find out the internal factors as well, partly by studying his body and partly by questioning him. The latter approach (or approaches, if you care to distinguish

the neurological and psychoanalytic approach to a man's interior) is much more time-consuming, leads to greater difficulties in forming theories, but is generally admitted to be the only hope for getting a complete theory.

As we have seen, there are three problems in predicting: getting all relevant facts, forming a complete theory, and carrying out the calculations (deductions) necessary for prediction. We have also seen that all of these run into serious difficulties even before we come to human thinking. Here the difficulties are much greater. It is practically impossible to get thorough knowledge of the internal structure of a man without killing him. He is such a complex machine (if he *is* a machine!) that our theories will have to be very complicated, and hence will take a very long time to discover. Finally, the problems of calculating results of his decisions are far beyond anything we have faced so far.

Some types of human actions are more easily predictable than others. There are cases where few facts are relevant and simple theories suffice. These are the actions classed roughly under the heading of reflexes or habits. We start with these, and as psychology progresses, we can predict more and more. There are also differences between people in this respect. We say that there are "unpredictable" people. These are the ones whose habits form a pattern too complex for us at the moment. There is no doubt that considerable progress will be made in this direction, but when the complexity of the problem is fully understood, there seems little chance of the psychologist ever being able to interfere seriously with our privacy of thinking.

To summarize: Complete predictability of events seems out of the question even in the inanimate world.

Prediction of human actions is perhaps the most complex problem facing the scientist. In this field we can expect only the very slowest progress, and we must expect that many types of decisions will forever remain outside the scientist's ability to predict.

Most important of all, we must remember that we are human beings. The very fact that a scientist observes us may alter our actions. This is the

greatest complicating factor for the psychologist, and it is also the simplest way to demonstrate the relation between predictions and free will. When we predict where a stone will fall, we may tell it just what it will do, and it will still do it. But tell a human being that he will use his right hand to scratch his ear, and you are likely to find that he uses it instead to punch you in the nose.

INTRODUCTION

Sperry directly addresses several issues in which the scientist and the humanist have often been at odds. In the first part of the article he looks at the importance of an understanding of conscious experience in order to understand the brain function. The basic problem here is the one that has been studied for centuries as the *mind-body problem*. When you read Sperry's statement that ". . . conscious mental forces do in fact govern and direct the nerve impulse traffic . . ." remember what Kemeny had to say about causal laws in general. The same question may be raised in reference to Sperry's statement that "it comes down to the issue of who pushes whom around in the population of causal forces that occupy the cranium."

In the central section of the article, Sperry gives his views on the question of free will. He wonders what would happen if individual decisions were completely "free" in the sense that they were totally unrelated to the individual's previous experience, emotional state, knowledge about relevant factors, and so on. To the extent that each of these factors enters in, we tend to think of an organism as having increasingly less "free will." Sperry's contribution is to call to our attention how very strange a person would be if he indeed were totally uninfluenced by all of these factors.

MIND, BRAIN, AND HUMANIST VALUES

Roger W. Sperry

Science, Antiscience, and Values

As a scientist invited to discuss humanist implications of the brain-behavior sciences, I find myself feeling a little like one who has been asked to mount the stand in self-defense as the accused. As they say back in Grade Two these days, for every action there is an equal and opposing reaction; and the recent sharp boom in science has not come without a corresponding rise in the voices of antiscience. Some of the going complaints in this regard are no doubt familiar: It is not only that science is going to blow us all off the globe, or crowd us off with its programs for death control, but that even the good things resulting from science—the sum total of all the better-things-for-better-living—have failed, we are told, to add substantially to a genuine satisfaction in living. And when it comes to the more profound humanist concerns, the reasons for living and the meaning and the value of it all, science seems only to take away and destroy, they say, and then refuses on principle to answer for its actions or even to be concerned with matters of values.

To some, even the objective explanatory progress that science is supposed to be making toward truth and the great central mystery of the universe begins to look like merely a handy system of humanoid guesses and correlational probabilities with no real verification possible. Others liken our explanatory progress to the penetration of a great maze that gets ever bleaker, the innermost chamber of which, should it ever be reached, being likely to hold exactly nothing or perhaps just the self-reflections of the scientists' own thought processes. And then, about as fast as our comprehension and control of nature goes up, antiscience sees man's rating in the grand design going down.

Before going on, I had better explain that the reference to values above and in the title was not accidental, though I well realize that any mixing of values and science tends to serve as a red flag in some quarters. Some of us may already be wondering, Since when do scientists presume to carry a license for discussions of values? Value judgments, we have all heard, lie outside the realm of science. Value matters are for popes and prophets, for philosophers and perhaps boy scout and YMCA leaders, but not for science or scientists. As a student of brain and behavior, I have

never been quite able to accept this. It seems the same as saying that value judgments lie outside the realm of knowledge and understanding. It is like saying that the best method we know of applying the human brain to problems of understanding must be discarded when it comes to problems of values. It is like saying that economics is riding under false colors in the National Science Foundation and ought to be exposed and expelled. And it is like saying that science is able to deal only with those phenomena and products of evolution that appeared prior to the emergence of higher brains, with their wants, needs, goal-directed properties, and, of course, the corresponding value systems that these impose.

Values have natural and logical origins. They are interdependent and interrelated in logical, hierarchical systems. These systems, and the perturbations thereof, ought to be subject to study and analysis and perhaps prediction and even some experimentation on a model basis these days, with computer assistance. I have always wondered whether rather little harm and perhaps much good in the long run might not come from opening to the free winds of scientific skepticism and inquiry even the most revered of our traditional and cultural values.

Humanist Impacts of Behavioral Science

We can now turn to our main question, What have been the major impacts, from the *humanist* standpoint, of the recent developments in the sciences that deal with mind and brain? At first glance the record achieved by the brain-behavior sciences during the past half-century

must seem to the humanist to read less like a list of contributions and advancements than like a list of major criminal offenses. The accusations that antiscience can raise in this area are not exactly trivial. For example, before science, man had reason to believe that he possessed a mind that was potent and full of something called "consciousness." Our modern experimental objective psychology and the neurosciences in general would divest the human brain of this fantasy and, in doing so, would dispense not only with the conscious mind but with most of the other spiritual components of human nature, including the immortal soul. Before science, man used to think that he was a spiritually free agent, possessing free will. Science tells us free will is just an illusion and gives us, instead, causal determinism. Where there used to be purpose and meaning in human behavior, science now shows us a complex biophysical machine with positive and negative feedback, composed entirely of material elements, all obeying the inexorable and universal laws of physics and chemistry. Thanks to Freud, with some assistance from astrophysics, science can be accused further of having deprived the thinking man of his Father in heaven, along with heaven itself. Freud's devastating indictment is said by many to have reduced much of man's formalized religion to little more than manifestations of neurosis.

Man's inner self and his heritage have not fared much better. Thanks to Darwin, and to Freud again, man now enters this life, not trailing clouds of glory, as the poet once had it, but trailing instead clouds of jungle-ism and bestiality, full of carnal impulses and a predisposition to Oedipal and other complexes. The confining veneer of our civilization is seen to be super-

ficial, and when it rubs thin or cracks, the animal within quickly shows through. These and related lesser onslaughts on the worth and the meaning of human nature tend to add up, one item reinforcing another, to yield a pretty dim over-all picture that is certainly not heartening to think about—and in science we generally don't think about it. Doubt and rejection of science by humanist thinkers in favor of other roads to truth is not hard to understand; and even for the scientist himself, the picture drawn by science imposes a severe test of his credo that it is better to know the truth, however ugly, and to live in accordance, than to live and die by false premises and illusory values.

But for myself, speaking as a brain researcher—and one not too familiar with matters ethical and philosophical and hence in a position to speak with some conviction—I find myself and my hypothetical working model for the brain to be in marked disagreement with many, if not the majority, of the foregoing implications especially and with that whole general picture of human nature that seems to emerge from the currently prevailing objective, materialistic approach of the brain-behavior sciences. When the humanist is led to favor the implications of modern materialism over the older idealistic values in these and related matters, I suspect that he has been taken, that science has sold society and itself a somewhat questionable bill of goods. There is not space here to present the whole story behind these remarks, and so I will try to concentrate selectively on what would seem to be the centermost issues, hoping that if the central foundation of the materialist view can be undermined the resultant crumbling in the upper structures will become evident.

The Nature of Consciousness: The Central Issue

Most of the disagreements that I have referred to revolve around, or hinge either directly or indirectly upon, a central point of controversy that emerges from the following question: Is it possible, in theory or in principle, to construct a complete, objective explanatory model of brain function without including consciousness in the causal sequence?

If the prevailing view in neuroscience is correct, that consciousness and mental forces in general can be ignored in our objective explanatory model, then we come out with materialism and all its implications. On the contrary, if it turns out that conscious mental forces do in fact govern and direct the nerve-impulse traffic and other biochemical and biophysical events in the brain and, hence, do have to be included as important features in the objective chain of control, then we come out at the opposite pole, or at mentalism, and with quite a different and more idealistic set of values all down the line. We deal here, of course, with the old mind-body dichotomy, the age-old problem of mind versus matter, the issue of the spiritual versus the material, on which books and philosophies have been written and philosophies have foundered ever since man started to think about his inner world and to question its relation to the outer "real" world.

Let us begin by stating the case against consciousness and mind as raised by today's objective experimental psychology, psychobiology, neurophysiology, and the related disciplines. The best way to deal with consciousness or introspective, subjective experience in any form, they tell

us, is to ignore it. Inner feelings and thoughts cannot be measured or weighed; they cannot be centrifuged or photographed, chromatographed, spectrographed, or otherwise recorded or dealt with objectively by any scientific methodology. As some kind of introspective, private, inner something, accessible only to the one experiencing individual, they simply must be excluded by policy from any scientific model or scientific explanation.

Furthermore, the neuroscientist of today feels he has a pretty fair idea about the kinds of things that excite and fire the nerve cells of the brain. Cell membrane changes, ion flow, chemical transmitters, pre- and post-synaptic potentials, sodium pump effects and the like, may be on his list of acceptable causal influences—but not consciousness. Consciousness, in the objective approach, is clearly made a second-rate citizen in the causal picture. It is relegated to the inferior status of (a) an inconsequential by-product, (b) an epiphenomenon (a sort of outsider on the inside), or most commonly, (c) just an inner aspect of the one material brain process. Scientists can see the brain as a complex, electrochemical communications network, full of nerve impulse traffic and other causally directed chemical and physical phenomena, with all elements moved by respectable scientific laws of physics, chemistry, physiology, and the like; but few are ready to tolerate an interjection into this causal machinery of any mental or conscious forces.

This is the general stance of modern behavioral science out of which comes today's prevailing objective, mechanistic, materialistic, behavioristic, fatalistic, reductionistic view of the nature of mind and psyche. This kind of thinking is not confined to our

laboratories and the classrooms, of course. It leaks and spreads, and though never officially imposed on the societies of the Western world, we nevertheless see the pervasive influence of creeping materialism everywhere we turn.

Once we have materialism squared off against mentalism in this way, I think we must all agree that neither is going to win the match on the basis of direct, factual evidence. The facts simply do not go far enough to provide the answer, or even to come close. Those centermost processes of the brain with which consciousness is presumably associated are simply not understood. They are so far beyond our comprehension at present that no one I know of has been able even to imagine their nature. We are speaking here of the brain code, the physiological language of the cerebral hemispheres. There is good reason to believe that this language is built of nerve impulses and related excitatory effects in nerve cells and fibers and perhaps also in those glia cells that are said to outnumber the nerve cells in the brain by about ten to one. And we would probably be safe in the further noncommittal statement that the brain code is built of spatiotemporal patterns of excitation. But when it comes to even imagining the critical variables in these patterns that correlate with the variables that we know in inner, conscious experience, we are still hopelessly lost.

Furthermore, the central unknowns directly associated with consciousness seem to be rather well cushioned on both the input and output sides of the brain by further zones of physiological unknowns. Our explanatory picture for brain function is reasonably satisfactory for the sensory input pathways and the distal portion of the

motor outflow. But that great in-between realm, starting at the stage where the incoming excitatory messages first reach the cortical surface of the brain, still today is very aptly referred to as the "mysterious black box."

To conclude that conscious, mental or psychic, forces have no place in filling this gap in our explanatory picture is at least to go well beyond the facts into the realm of intuition and speculation. The objective, materialist doctrine of behavioral science, which tends to be identified with a rigorous scientific approach, is thus seen to rest, in fact, on an insupportable mental inference that goes far beyond the objective evidence and hence is founded on the cardinal sin of science. One can still find here and there in the literature a modicum of some final, perhaps "last rite," respect paid to the psyche. For example, there is the acceptance by Charles Sherrington of the possible coexistence of two separate phenomenal realms in the brain, and there is the stand of Carl Rogers that man's inner experience must be recognized as well as the brain mechanism of objective psychology. In the existence of two such very different realms, Rogers sees a lasting paradox with which we all must learn to live. But even the dualists are quite prepared to go along these days with the conviction held by most brain researchers—up to some 99.9 per cent of us, I suppose—that conscious mental forces can be safely ignored, insofar as the objective, scientific study of the brain is concerned.

An Alternative Mentalist Position

In the pages that follow, I am going to line myself up with the 0.1 per cent or so mentalist minority in a stand

that admittedly also goes well beyond the facts. It is a position, however, that seems to me equally strong and somewhat more appealing than those we have just outlined. In my own hypothetical brain model, conscious awareness does get representation as a very real causal agent and rates an important place in the causal sequence and chain of control in brain events, in which it appears as an active, operational force. Any model or description that leaves out conscious forces, according to this view, is bound to be sadly incomplete and unsatisfactory. The conscious mind in this scheme, far from being put aside as a by-product, epiphenomenon, or inner aspect, is located front and center, directly in the midst of the causal interplay of cerebral mechanisms. Mind and consciousness are put in the driver's seat, as it were: They give the orders, and they push and haul around the physiology and the physical and chemical processes as much as or more than the latter processes direct them. This scheme is one that puts mind back over matter, in a sense, not under or outside or beside it. It is a scheme that idealizes ideas and ideals over physical and chemical interactions, nerve impulse traffic, and DNA. It is a brain model in which conscious mental psychic forces are recognized to be the crowning achievement of some five hundred million years or more of evolution.

Now, what is the argument in favor of mentalism, the argument that holds that ideas and other mental entities push around the physiological and biochemical events in the brain? The argument is simple and goes as follows: First, it contends that mind and consciousness are dynamic, emergent (pattern or configurational) properties of the living brain in action. There are

usually plenty of "takers" on this first point, including even some of the tough-minded brain researchers, as, for example, the outstanding neuroanatomist, C. J. Herrick. Second, the argument goes a critical step farther and insists that these emergent properties in the brain have causal potency—just as they do elsewhere in the universe. And there we have the simple answer to the age-old enigma of consciousness. Who was it who said, that nothing is so simple as yesterday's solution, nothing so complicated as tomorrow's problems?

But let us spell out this answer a little further, since this whole subject has at times been a bit complicated. To put it very simply, it comes down to the issue of who pushes whom around in the population of causal forces that occupy the cranium. It is a matter, in other words, of straightening out the peck-order hierarchy among intracranial control agents. There exists within the cranium a whole world of diverse causal forces; what is more, there are forces within forces within forces, as in no other cubic half-foot of universe that we know. At the lowermost levels in this system, we have local aggregates of some sixty or more types of subnuclear particles interacting with great energy, all within the neutrons and protons of their respective atomic nuclei. These chaps, of course, do not have very much to say about what goes on in the affairs of the brain. We can pretty well forget them, because they are all firmly trapped and kept in line by their atomic overseers. The atomic nuclei and associated electrons are also, of course, firmly controlled in turn. The various atomic elements are "molecule-bound"—that is, they are hauled and pushed around by the larger spatial and configurational

forces of their encompassing molecules.

Similarly, the molecules of the brain are themselves pretty well bound up and ordered around by their respective cells and tissues. Along with all of their internal atomic and subnuclear parts and their neighboring molecular partners, the brain molecules are obliged to submit to a course of activity in time and space that is very largely determined, for the lifetime of any given cell, by the over-all dynamic and spatial properties of the whole cell as an entity. Even the brain cells, however, with their long fibers and impulse-conducting properties, do not have very much to say about when they are going to fire their messages, for example, or in what time pattern they will fire them. The firing orders for the day come from a higher command.

In other words, the flow and the timing of impulse traffic through any brain cell, or even a nucleus of cells in the brain, are governed largely by the over-all encompassing properties of the whole cerebral circuit system, within which the given cells and fibers are incorporated, and also by the relationship of this circuit system to other circuit systems. Further, the dynamic properties of the cerebral system as a whole, and the way in which these properties direct and govern the flow of impulse traffic throughout the system—that is, the general circuit properties of the whole brain—may undergo radical and widespread changes from one moment to the next with just the flick of a cerebral facilitatory "set." This set is a shifting pattern of central excitation that will open or prime one group of circuit pathways with its own special pattern properties, while at the same time closing, repressing, or inhibiting endless other circuit poten-

tialities that might otherwise be open and available for impulse traffic. These changes of "set" are responsible, for example, for such things as a shift of attention, a turn of thought, a change of feeling, or a new insight. To make a long story short, if one keeps climbing upward in the chain of command within the brain, one finds at the very top those over-all organizational forces and dynamic properties of the large patterns of cerebral excitation that are correlated with mental states or psychic activity. And this brings us close to the main issue.

We can take this argument a step further by looking at an illustrative example of one of these mental entities. For simplicity, let us consider an elemental sensation. Instead of philosophy's old favorite, the color red (the philosophic and geographic locus of which seems sometimes to be in some doubt), let us use another example, pain. To be more specific, let us say we are talking about pain in the fingers and thumb of the left hand, and let us pin it down further to pain in the left hand of an arm that was amputated above the elbow some months previously. You will recall that the suffering caused by pain localized mentally in a phantom limb is no easier to bear than that in a limb that is still there. It will be easier, however, by using this example, for us to infer where our conscious awareness does not reside.

In regard to the pain in a phantom limb, my contention is that any groans it may elicit from our patient and any other response measures or behavioral outputs that may be taken to be the result of the pain sensation are indeed caused not by the biophysics, chemistry, or physiology of the cerebral nerve impulses as such, but by the pain quality, the pain property, *per se*. This brings us, then, to the real crux of

the argument. Nerve excitations are just as common to pleasure, of course, as to pain, and the same is true of any other sensation. What is critical is that unique patterning of cerebral excitation that produces pain instead of something else. It is the over-all functional property of this pain pattern as a pattern that is critical in the causal sequence of brain affairs. This pattern has a dynamic entity, the qualitative effect of which must be conceived functionally and operationally and in terms of its impact on a living, unanesthetized cerebral system. It is this over-all pattern effect in brain dynamics that is the pain quality of inner experience. To try to explain the pain pattern or any other mental qualities only in terms of the spatiotemporal arrangement of nerve impulses, without reference to the mental properties and the mental qualities themselves, would be as formidable as trying to describe any of the endless variety of complex molecular reactions known to biochemistry wholly in terms of the properties of the electron, proton, and neutron and their subnuclear particles plus (and this, of course, is critical) their spatiotemporal relationships. By including the spatiotemporal relations, such a description becomes feasible in theory, probably, but fantastically impractical. Moreover, by the time science arrives at a point where it can describe the critical details of the impulse pattern of a mental experience in the functional terms and setting required, it will be describing, in effect, the conscious force or property itself. When we reach such a point, the conscious force will be recognized as such, and we shall be calling it just that—or at least that is the hypothesis I am putting forward.

Many readers will note my reliance throughout this discussion on the

emergent concepts of T. H. Morgan and the corresponding configurational and field concepts of Gestalt psychology. The Gestalt schools of psychology and philosophy went wrong only when they moved into the brain and tried to transfer their pattern properties directly from the outside world and sensory surfaces into the cerebral cortex. The central, emergent conscious force within the brain, as visualized here, is not a simple surrounding envelope, or volume property, or any other kind of "isomorph," as the Gestalt schools tried to make it. It is rather a functional pattern that has to be worked out in entirely new terms, that is, in terms of the functional circuitry of the brain, in terms of the still unknown brain code.

Above simple pain and other sensations in brain dynamics, we find, of course, the more complex but equally potent forces of perception, emotion, reason, belief, insight, judgment, cognition, and all the rest. In the onward flow of conscious brain states, one state calling up the next, these are the kinds of dynamic entities that call the plays. It is exactly these encompassing mental forces that direct and govern the inner impulse traffic, including its electrochemical and biophysical aspects. When trying to visualize mental properties as they have been described, it is important to keep in mind the fact that all of the simpler, more primitive, electric, atomic, molecular, cellular, and physiological forces remain present, of course, and they all continue to operate. None has been cancelled; but these lower level forces and properties have been superseded, encompassed, as it were, by those forces of successively higher organizational entities. We must remember in particular that, for the transmission of nerve impulses, all of the usual electri-

cal, chemical, and physiological laws still apply at the level of the cell, the fiber, and the synaptic junction. We must remember further that proper function in the uppermost levels always depends on normal operation at subsidiary levels.

Near the apex of this command system in the brain—to return to more humanistic concerns—we find ideas. Man over the chimpanzee has ideas and ideals. In the brain model proposed here, the causal potency of an idea, or an ideal, becomes just as real as that of a molecule, a cell, or a nerve impulse. Ideas cause ideas and help evolve new ideas. They interact with each other and with other mental forces in the same brain, in neighboring brains, and thanks to global communication, in far distant, foreign brains. And they also interact with the external surroundings to produce in toto a burstwise advance in evolution that is far beyond anything to hit the evolutionary scene yet, including the emergence of the living cell.

In the proposed scheme, the interplay of psychic and mental forces, though accessible—like the interior of the earth—only indirectly at this date becomes, in principle, a proper phenomenon for scientific investigation. Aside from problems of complexity and adequate technology, there would seem to be no great obstacle in principle to the eventual objective, scientific treatment of mental phenomena. One may see statements in the literature these days discouraging the hope that the mind is capable of explaining itself in terms of its own ideas; the argument is that no machine, living or otherwise, can logically embody within itself a complete description of itself. When you read such statements, however, always underline that word "complete" and then consider the extent of

the explanatory possibilities that still remain even though they fall somewhat short of complete. Underline also that word "itself" and remember that this kind of logic does not prevent a man's mind from acquiring a complete description of his neighbor's mind or from passing on this description to other neighbors, excepting only the one he has described.

For an outside, second, brain, however, to directly experience the subjective qualities in an observed brain, it would seem to be necessary for the detector brain in the observer to be coupled in parallel to the emitting brain and wired directly into the specialized cerebral circuitry involved. This does not look very feasible under ordinary conditions for the near future. However, we do seem to be approaching exactly this situation experimentally in recent studies in which the brains of cats and monkeys have been bisected down the midplane into right and left halves. In the surgical process, we may leave a few cross-connections, coupling selected cerebral centers between "mind-right" and "mind-left." When the midline disconnection is complete, two separate mentalities are the result, which sense, perceive, learn, and remember independently. Each half seems to have its own realm of conscious awareness, and each is apparently as much out of contact with the inner experience of the other as are two brains in separate skulls. But when a band of cross-connections is left intact, linking, for example, the right and left centers for vision or those for touch sensibility in the hands, the inner, mental, subjective experience of the one brain seems to become available to the other.

Something of the kind can also be seen in studies of human patients who have had a similar surgical disconnec-

tion of the hemispheres for medical or therapeutic purposes and in whom cross-connections have been left intact between the lower brain centers involved in emotion and feeling. Whereas the cognitive, perceptual, mnemonic, and related experiences of mind-right in these people seem to be entirely out of touch with the corresponding experiences of mind-left, each brain half seems to share the emotional experience of the other. For example, if an emotion is triggered through vision by the introduction of an unexpected pin-up picture of a nude into a sequence of ordinary geometric pattern stimuli being projected into only one brain half, it is quite apparent from the verbal read-out through the other half of the brain (that is, the one not directly excited) that this second hemisphere also feels properly embarrassed—or whatever the case may be. The second hemisphere, however, has no idea why it has these inner feelings and is unable to describe their source.

Looking back from this point, you will note that the earlier basic distinction or dichotomy between mentalism and materialism is resolved in this interpretation, and the former polar differences with respect to human values, when recast in the present scheme, become mainly errors of reductionism. This may be easily recognized as the old "nothing but" fallacy; that is, the tendency, in the present case, to reduce mind to nothing but brain mechanism, or thought to nothing but a flow of nerve impulses. For those acquainted with theories of mind, the new twist here, if any, is to be found in the attempt to make the emergent properties of inner experience conform to the inner brain code, rather than to the outside world or subjective impressions or sensory

patterns; combined, of course, with the critical interjection of these mental qualities into the causal sequence. Note that we have not rejected the objective approach of science; it is an objective explanatory model that we are discussing. Our quarrel is not with the objective approach but with the long accepted demand for exclusion of mental forces, psychic properties, and conscious qualities—what the physicist might class as “higher-order effects” or “co-operative effects”—from the objective scientific explanation.

The present scheme would put mind back into the brain of objective science and in a position of top command. If correct, it would eliminate the old dualistic confusions, the dichotomies and the paradoxes, proposing instead a single unified system extending from subnuclear forces near the bottom up through ideas at the top. As a scientific theory of mind, it would provide a long sought unifying view on which to base our conception of human nature, the kind of view whose lack has recently been deplored in leading articles in *Science* and elsewhere. Moreover, this scheme suggests a possible answer not only for the relation between mind and brain but also for that between the outside world and its inner cerebral representation, another conundrum since the days of Plato. When used as a conceptual skeleton on which to build a body of philosophy, this theory tends to favor somewhat a single “this world” measuring stick for evaluating man and existence. As for the older materialist doctrine, one can say, in summary, that the denial or downgrading of conscious mental forces in objective experimental psychology during the past half-century has been tremendously valuable and successful as a tactical

expedient for a developing science. It is hardly a doctrine, however, on which to build societal philosophy and cultural values.

Free Will

Let us shift gears at this point and consider another outcome of the mind-brain sciences that appears to run a close second in its threat to cherished images of human nature. This humanist “Enemy Number 2” to which I refer (some would rate it Number 1) is the scientific rejection of free will. Every advance in the science of behavior, whether it has come from the psychiatrist’s couch, from microelectrode recording, or from the use of psychotropic drugs, brain splitting, Skinner boxes or the electron microscope, seems only to reinforce that old suspicion that free will is just an illusion. Like most others in brain research, I work on the assumption that every apparently free mental choice that I or anyone else has ever made must in fact have been causally predetermined in the preceding brain states and related events. This means that any decision in which any of us has ever engaged could not possibly have had any other outcome. It means none of us, hearing or reading these words, had any real chance to be doing anything else at this particular moment. It means that we are now and always have been imprisoned, as it were, in the inexorable onward march of causal determinism.

Attempts to restore free will to the human brain by recourse to various forms of indeterminacy—physical, logical, emergent, or otherwise—have failed so far as I can see to do more than introduce a bit of unpredictable caprice into our comportment that most of us would prefer to be without.

Neither science nor philosophy seems able to refute the old admonition that "the moving finger writes; and, having writ, moves on." And piety, wit, and tears still seem impotent to change this situation. I do not feel overly comfortable about this kind of thinking any more than anyone else, but as yet I have not found any way out of it.

But before we start drawing gloomy humanist deductions from this apparent inevitability, concluding that moral responsibility is thereby removed or, on the other hand, simply rejecting science and determinism on emotional grounds, we should bear in mind a few additional points. In the present scheme, these points add up to the conclusion that if we really did have freedom of choice we might very likely prefer not to have; that is, we might well prefer to leave determinism in control, exactly as science postulates. It should be clear by now that in the brain model described here, man is provided in large measure with the mental forces and the mental ability to determine his own actions. This scheme thus allows a high degree of freedom from outside forces as well as mastery over the inner cellular, molecular, and atomic aspects of brain activity. Depending on the state of one's will power, the model also allows considerable freedom from lower-level natural impulses and even from occasional thoughts, beliefs, and the like, though not, of course, from the whole complex. In other words, the kind of brain visualized here does indeed give man plenty of free will, provided we think of free will as self-determination. To a very real and large extent, a person does determine with his own mind what he is going to do from among a large number of possibilities. This does not mean, however, that he

is free from the forces of his own decision-making machinery. In particular, what this present model does not do is to free a person from the combined effects of his own thoughts, his own reasoning, his own feeling, his own beliefs, ideals, and hopes, nor does it free him from his inherited makeup or his lifetime memories. All these and more, including, yes, unconscious desires, exert in the brain their due causal influence upon any mental decision, and the combined resultant determines the inevitable but self-determined and highly special and highly personal outcome. We thus come back to the question, Do we really want free will, in the sense of gaining freedom from our own minds, from our own selves, and hence, from everything most precious that makes us us?

There is a bit more to the story of how one may "learn to stop worrying about freedom and come to love determinism," but it boils down to the old saying, "If you can't lick 'em, join 'em." Or as Confucius might have said, "If fate inevitable, relax and enjoy." Or to put it more directly, "There may be worse fates than causal determinism." Maybe after all it is better to be imbedded firmly in the causal flow of cosmic forces, as an integral part thereof, than to be on the loose and out of contact with these forces—free-floating as it were—with behavioral possibilities that have no antecedent cause and hence no reason or any reliability when it comes to future plans, predictions, or promises. I suspect, in fact, that if you were assigned the task of trying to design and build the perfect free-will model (let us say the perfect all-wise decision-making machine to top all competitors' decision-making machines), it is possible that you might decide that

your aim should not be so much to free the machinery from causal contact as the opposite; that is, to try to incorporate into your model the potential value of universal causal contact—in other words, contact with all related information in proper proportion, past, present, and future.

In any case, it is clear that the human brain has come a long way in evolution in exactly this direction when you consider the amount and the kind of causal factors that this multidimensional, intracranial vortex draws into itself, scans, and brings out bear on the process of turning out one of its pre-ordained decisions. Potentially included, thanks to memory, are the events and collected wisdom of most of a human lifetime. We can also include, potentially, given a trip to the library, the accumulated knowledge of all recorded history. And we must add to all the foregoing, thanks to reason and logic, much of the future forecast and predictive value extractable for all this data. Maybe the total falls a bit short of universal causal contact, maybe it's not quite up to the kind of thing that evolution has going for it over on Galaxy Nine, and maybe, in spite of all, any decision that comes out is still predetermined. Nevertheless, it still represents a very long jump in the direction of freedom from the primeval slime mold, the Jurassic sand dollar, or even the latest 1965 model Orangutan.¹

We may note in passing that on the debit side of the ledger there is little in our proposed model for conscious-

ness to bolster one's hopes either for extrasensory perception or for post-mortem perception. Similarly, pre-partum perception in the embryo may be presumed not to amount to much until after the requisite cerebral machinery for conscious awareness has begun to attain functional maturity in the later months of fetal life and in subsequent postnatal development.

Plasticity of Human Nature and Inheritance of Behavior Traits

Finally, in connection with development, I must mention briefly certain other advances in the brain-behavior sciences that have resulted in important revisions during the past two decades in our general conception of human nature. These advances have concerned the extent to which behavior traits can be inherited and the extent to which human nature is plastic and subject to shaping by experience and environment.

Through most of the first part of this century and up until about twenty years ago, the view prevailed that the brain gets its start in fetal life as an essentially equipotential network, functionally unstructured, a blank slate, as it were, which is then gradually channelized from early fetal movements onward by functional trial and error, practice, conditioning, learning, and experience. The objective, materialist movement in psychology, which was established first in Russia, largely under the influence of Pavlov, and which appeared soon afterward in this country, pioneered by Watson, under the name "behaviorism," has been identified almost as much with the promotion and idolatry of the conditioned response as it has with the demotion and vilification of consciousness. In this doctrine the

¹This paragraph and the preceding one are taken almost verbatim from an earlier article, *Problems Outstanding in the Evolution of Brain Function: James Arthur Lecture of the Evolution of the Human Brain* (New York: American Museum of Natural History, 1964).

mind, or psyche, was believed to develop gradually out of a life-long chain of successive conditioned-reflex associations, starting in the infant from a few elementary reactions, like love and hate, fear and anger. The whole idea of the genetic inheritance of behavior patterns was forcibly renounced, until the term "instinct" became highly discredited in professional circles, its defamation almost equalling that of consciousness. In those days, the embryonic growth of brain pathways was believed to be by nature non-selective and diffuse, and the establishment of precise fiber connections was held to be unimportant anyway for orderly function. The nerve connections, once laid down, were thought to be able to undergo radical and wholesale rearrangement by surgery, injury, and regeneration without disrupting orderly function. In the scientific thinking of those times, the brain was endowed with an almost mysteriously omnipotent plasticity and readaptation capacity. In general, science seemed to be telling us through the twenties and thirties and into the early forties that the human brain and human nature as well were extreme in their malleability. It seemed at that time a scientifically sound conclusion that it would be possible, by means of an appropriate program of training and environmental conditioning, to shape human nature, and hence society, within wide limits into a desired mold.

Much of the basic scientific thinking and evidence behind this view has since suffered a series of severe upsets, leading to a current stand that is almost diametrically opposed to the earlier doctrines. Instead of a loose, universal plasticity in brain hookups, we now see a basic built-in wiring diagram, characteristic of the species

and functionally rather rigid. Instead of diffuse, non-selective growth of nerve connections in brain development, we now see a very precise and highly ordered patterning of brain fiber pathways and connections, all strictly preregulated by specific genetic effects and cytochemical affinities. Where there used to be an outright denunciation of the whole concept of "instinct," we now accept the idea that an entire evolutionary tree can be worked out on the basis of inherited behavior patterns, just as it can be worked out on the basis of morphological or serological traits. The conditioned response, along with other forms of learning, continues to be recognized, of course, as a highly powerful modeling influence, especially in man, but only within limits much narrower than previously supposed.

Within the specialized fields of scientific inquiry involved here, the pendulum of opinion continues at this date to swing in the direction of inheritance. How far it will go can only be guessed. It is still too soon for the implications to have fully permeated even the neighboring scientific disciplines. What implications these changes in the basic brain-behavior sciences may have, if any, for more distant problems in the social sciences will take much longer to evaluate. The latter, of course, will have to be worked out in their own right and at their own level. In any case, it would seem that the evidence available today says that we should renounce, along with other aspects of the behaviorist, materialist approach discussed above, the old Pavlovian-Watsonian conditioned-reflex theory of the psyche, with its radical environmentalism that used to tell us that literally 99 per cent of human nature and mind

is a product of experience and training.

Our re-examination of the materialist doctrine in psychology could be extended much further into matters far removed from those in which brain researchers feel at all comfortable. Let me only remind my readers that the peck-order of causal entities does not stop within the individual brain but goes on up into higher levels involving society and culture, various subentities of which must be properly credited with many of man's most remarkable and fantastic achievements.

Reference to society brings, of course, the pressing reminder that any attempt to upgrade human nature through a more idealistic conception of mind is bound these days to be overwhelmingly counteracted by the cold laws of mathematics and the devastating downgrading effect of surplus numbers on the worth of the individual. We do not need the third law of psychodynamics to tell us that the optimum carrying capacity of our globe is perhaps already exceeded from the standpoint of quality, dignity, meaning, and value for the human individual. When we look at the rising threat posed by the effects of human surplus and its by-products on the hard-won and painstaking achievements of eons of evolution, we are inclined to forget our little ideological skirmish with materialism, along with most of the human betterment efforts of our times, as just another losing battle in the face of mounting humanity—effort down the drain, until some higher force in our mental hierarchy than natural impulse can be brought to bear.

When it comes to the future outlook and an attempt to make predictions regarding the future of man, behavioral science is hampered by a

technical difficulty, in that once the prediction is published and man becomes acquainted with what he is supposed to do, he is in a position to take the prediction into account and is apt to be just perverse enough to do the reverse. Keeping this in mind, we can forecast that our generation and future generations need not really worry about surplus numbers, or who will outbreed whom, or any of the other problems we have touched on above, because these and related matters promise to be settled shortly in that final fatal flare of fission fireworks.

But to return to the central concern of this essay—the impact of creeping materialism in the brain-behavior sciences—we can say in summary that it is possible to see today an objective, explanatory model of brain function that neither contradicts nor degrades but rather affirms age-old humanist values, ideals, and meaning in human endeavor. The noble, free, or heroic, the exalted or sublime, qualities—or the opposite, for that is how meanings arise—that the humanist formerly thought he could see in man and his activities are present in our model, much as history and common experience have always shown. Finally, for those who like to receive a take-home lesson, that of the foregoing is simple for scientist and humanist alike: Never underestimate the power of an ideal.

Related Articles by the Author

- "Embryogenesis of Behavioral Nerve Nets," in *Organogenesis*, ed. R. L. DeHaan and H. Ursprung (New York: Holt, Rinehart & Winston, Inc., 1965).
- "The Great Cerebral Commissure," *Scientific American*, CCX (1964), 42–52.

this common conviction, on the one hand with essentially mechanistic views of the workings of our brains and the universe, or on the other with predestinarian views of human history and destiny. If any standpoint exists from which all I do can be seen as part of a single, coherent and determinate structure in space-time, then surely my sense of freedom in facing a decision, no matter who shares it, must be dismissed as an illusion, reflecting only my ignorance of the true facts at the time?

In view of Eddington's long preoccupation with the deepest implications of physical theory, it is not surprising to find recurrent references to this ancient dilemma in his philosophical writings. 'In the old conflict between freewill and predestination', he wrote in his famous Gifford Lectures of 1927, 'it has seemed hitherto that physics comes down heavily on the side of predestination. Without making extravagant claims for the scope of natural law, its moral sympathy has been with the view that whatever the future may bring forth is already foretold in the configurations of the past:

'Yea, the first Morning of Creation
wrote
What the Last Dawn of Reckoning
shall read.'

'I am not so rash', he went on, 'as to invade Scotland with a solution of a problem which has rent her from the synod to the cottage'; but he did proceed, then and later, to explore with characteristic vigour and imagination the possibilities left open, as he saw it, by the advent in physics of what became known as Heisenberg's Principle of Uncertainty, or, as Ed-

dington preferred to call it, the Principle of Indeterminacy. 'The future is a combination of the causal influences of the past together with unpredictable elements—unpredictable not merely because it is impracticable to obtain the data of prediction, but because no data connected causally with our experience exist. . . . Science', he continued, 'thereby withdraws its moral opposition to freewill.'¹

In his *Turner Lectures* eleven years later (1938) Eddington was careful to reject the inference that freedom of will amounted only to the operation of chance factors in the brain. On the contrary, he argued ingeniously for a theory of volition as imposing *correlations* on the atomic motions left undetermined by physical theory, so that although physically unpredictable they could be both law-abiding and significant in their effects on a molar scale.²

2. A Prior Question

But the details of Eddington's theory of 'conscious matter', even if I felt competent to expound or criticize them, are not here our main concern. As a physicist who is now engaged in brain research, I am peacefully committed to the working assumption that brain-processes should be expected to follow normal physical laws', like any other part of the material universe, until proved to do otherwise. The question I wish to raise in a sense undercuts the debate in which Eddington found himself engaged, for although I am unequivocally on his side against those who debunk human freewill, I believe that the fallacy in their argument enters a step farther back—a step which indeed Eddington himself, like most

disputants in this field, seems to have been prepared to take for granted. How cogent, in fact, is the initial assumption that a fully mechanistic theory of our brain-processes would prove our actions inevitable, and our freedom an illusion?

At the outset, I want to shake loose from a whole family of arguments based on the limits of our power to observe the brain and to make predictions from what we observe. These limits are real enough, and relevant enough to many a dispute over what is likely to be possible *in practice* by way of the prediction and control of human action; but I suggest that they are both inadequate and unnecessary as bulwarks against the attack on human freedom that concerns us here. In fact, as we shall see, a defence based on limits of this kind can be dangerously misleading, by deflecting attention from the logical error that I believe vitiates the original attack.

So to clarify our ideas I suggest that we start with the most extreme case, and ask ourselves what would follow if mechanistic brain theory were to develop, beyond the wildest dreams of any of my colleagues in brain physiology, into a fully deterministic science. Suppose that all the relevant facts of the workings of your brain could be made available, without disturbing it, to a computer system capable of predicting its future behaviour from these facts and the environmental forces acting on your nervous system. Suppose, in other words, that your brain were as mechanical as clockwork, and as accessible to deterministic analysis. What then?

I propose to argue that even then—even with this wildly idealized pre-Heisenberg conception of a physically

determinate brain—the denial of human freedom in general would be not only unfounded but demonstrably mistaken. Lest this seem too much to swallow, let me say that the mistake in question results from neglecting a small factor which in almost every other context can safely be neglected: namely, the change that must take place in the physical brain (according to mechanistic brain theory itself) when knowledge is acquired by a human agent. The conclusions to which we are led by taking account of this factor are I think inescapable, and demand no equipment other than meticulous patience; but once grasped they are far-reaching in their implications.

3. The Costs of Human Knowledge

Let us begin by reminding ourselves that in science, as in every intellectual enterprise, the ultimate object is to reach *valid beliefs*. We may be tempted to equate this with the enunciation of *true propositions*; but here, for reasons which will become apparent, the temptation must be resisted. Particularly when the scientific spotlight is turned upon the process of cerebration itself, we can no longer afford to confuse the validity of belief with the truth of propositions; for if a proposition happens to describe, implicitly or explicitly, the brain-state which reflects what you believe, *its accuracy may depend upon whether or not you believe it*.

Here we come to the core of our later discussion. We are familiar—thanks in no small measure to the popular writing of Eddington—with the general idea that scientific knowledge has certain irreducible physical costs. The process of observation in

fact has three stages, in which information is respectively (i) generated, (ii) transmitted, and finally (iii) absorbed. At each we have to pay for our information by accepting limitations which are often negligible, but may assume crucial importance in special cases.

(i) *Our information about the physical world is generated only at the cost of interaction, between some 'probe' and what we are observing*

The amount of interaction is not reducible without limit, but must exceed a quantum (defined by Planck's Constant h) which is no longer negligible when the process under observation is itself one of small-scale quantum exchange. Hence follows (eventually) the Heisenberg Principle, according to which no observation can yield a picture which is both completely precise and completely up to date; the more information we generate, the more unpredictably disturbed we leave the system by our observations.

(ii) *Information takes time to be transmitted to us*

The time required is not reducible without limit, but must exceed an interval x/c , where x is the distance to be traversed, and c is the velocity of light. This interval is no longer negligible when the process under observation involves velocities comparable with that of light. Hence follows (eventually) Einstein's Special Relativity Principle, whereby two observers moving with different velocities must in general *differ* in some of their conclusions if both are observing accurately—and neither can claim to possess the 'one true picture'.

4. Logical Indeterminacy

But the cost to which I want to draw attention arises, as one consequence of the mechanistic brain-theory itself, at the third stage.

(iii) *Information can be absorbed only at the cost of changes in the brain*

The amount of physical change required is not reducible without limit, for if we suppose that what a person believes is rigorously represented by the state of what we may call his cognitive mechanism, then discrete, non-negligible changes in what he believes must entail discrete, non-negligible changes in that state.

It follows that even if the brain were as mechanical as clockwork, no completely detailed present or future¹⁴ description of a man's brain can be equally accurate whether the man believes it or not. (a) It may be accurate *before* he believes it, and then it would automatically be rendered out of date by the brain-changes produced by his believing it; or (b) it might be possible to arrange that the brain-changes produced by his believing it would bring his brain into the state it describes, in which case it must be inaccurate *unless* he believes it, so he would not be in error to *disbelieve* it.

In either case, then, the brain-description lacks the 'take it or leave it' character of scientific descriptions of the rest of the physical world, since its validity depends precisely upon whether the subject takes it or leaves it! True, any number of *detached* observers could predict whether the subject will 'take it' or 'leave it'; but this prediction in turn, though valid for

the observers in detachment, would lack any 'take it or leave it' validity for the subject. It would still be true that for such brain-states, and future events causally dependent upon them, no *universally valid (pre)determination exists*: no complete and certain prediction waits undiscovered upon which the subject and his observers would be correct to agree.

Notice that we are not saying only that the subject cannot *make* or *discover* a prediction of his future brain-states, but that there *exists* no definitive prediction that could claim his assent. There is thus no question here of 'ignorance of the truth' on the subject's part, nor are we concerned with what he subjectively *feels* about the future event. What we are saying is that as a matter of objective fact, whatever he feels about it, the event in question is one about which for him *no definitive information yet exists*.

It would be nonsensical, and not just ineffectual, for him to 'wish he knew' what form the event would take; for on that matter there is nothing final for him to know. Even though the whole physical process concerned were as mechanical as clockwork, his special relation to it makes an observer's view for him not just unknowable but logically invalid and inconclusive. For him, strange as it may seem, the only logically admissible view is that this particular future event is as yet *indeterminate*.

The foregoing 'non-existence theorem' I have elsewhere³ called the Principle of Logical Indeterminacy—perhaps not the best name for it, but designed to distinguish it from the Principle of Physical Indeterminacy enunciated by Heisenberg. It applies, as we have seen, even in a physically determinate universe; but the admis-

sion of physical indeterminacy to our postulated brain theory would clearly do nothing to diminish its force.

5. People in Dialogue

It is easy to see that this logical indeterminacy must apply not only to present or future processes in the subject's own brain and their causal consequences, but also to those in the brains of other people with whom he is in dialogue, or with whom dialogue is possible.⁴ From a physical point of view, the two-way interaction between people in dialogue (as distinct from purely manipulative monologue) introduces a coupling between the physical states of their cognitive mechanisms. They thus become effectively one system for purposes of mechanistic analysis, so that to this joint cognitive system, no matter how many people are involved, the foregoing argument again applies. No party to a dialogue can logically regard any of the others as a fully determinate system. We are not merely saying that nobody in dialogue *should* or *ought* to do so, nor that it is impractical to do so, but that it is *contrary to the mechanistic brain theory itself*. In this technical sense, people in dialogue are irreducibly *mysterious* to one another as well as to themselves.

Here—if not before—someone may suspect a trick. We have claimed to be working out some consequences of the most crassly mechanistic view of the brain, as a clockwork mechanism within a clockwork universe. How on earth have we moved from such a starting point to talking of people as irreducibly indeterminate and mysterious to one another? Something, surely, has been surreptitiously imported along the way?

And of course something crucial has been imported, though not surreptitiously. This is the assumption made by the mechanist himself, that what a man believes and thinks is rigorously reflected in the state of his brain. It is this very assumption—one that to many antimechanists has appeared as a deadly challenge—which generates the peculiar logical consequences. It amounts to postulating that certain parts of the domain of physical science—namely, the brains of cognitive subjects—have the task of *representing that domain of discourse itself*. Since sentences that talk about themselves are notoriously treacherous in logic, we need not be surprised that we have to accept some indeterminacy—some lack of complete detail—in any description of brain machinery that must be self-consistently representable by that same brain machinery.*

No, the root of mystery in our analysis is the brute fact of cognition itself. Deny that the individuals under discussion are cognitive subjects, and the question of what they would be correct to believe disappears. The whole physical domain of discourse would thus become fully specifiable in theory—but then, of course, we would no longer be talking about human beings like ourselves, and about what our brain science allows us to believe, which was the original purpose of our discussion. In a sense, then,

* The basic logical point here has been cogently argued in connection with the limitations of predictors by Sir Karl Popper, in a well-known paper on Indeterminism in Classical and Quantum Physics,⁵ where he shows that for this reason no computing machine, however 'deterministic' in its workings, could be used to calculate its own future in every detail.

all we have been doing here is to spell out some implications of the mysterious but empirically undeniable correlation between our conscious experience and what goes on inside our heads.

6. Freedom of Action

Let us now see how what we call a free choice appears in the light of our discussion. Suppose you have to decide at 7 p.m. whether to leave for the 7.15 p.m. train, or to wait for a later one. This means that just before 7 p.m. there are two possible descriptions of the future event at 7 p.m.: either (G) you decide to go, or (W) you decide to wait. Until 7 p.m. you contemplate G and W as *undetermined* alternatives. At 7 p.m. you decide to go, thus making description G true and W false; and you act freely as far as you can tell. But next day a Determinist calls with a ciné film of data taken from your brain before 7 p.m., from which he can prove that the outcome was calculable in advance to be G. Does this evidence retrospectively refute your belief that your action was free?

If 'free' were defined to mean 'unpredictable by anyone,' then of course in that sense your belief would be falsified. I suggest, however, that this definition, though true perhaps of the 'freedom' of *caprice*, begs too many questions to be adopted uncritically for the 'freedom' of responsible action. Indeed on further analysis it seems to put the emphasis in quite the wrong place. An action might be rendered totally unpredictable by some random physical disturbance of your neural mechanism; but, as Eddington among others pointed out, this could well be

held to diminish rather than enhance your responsibility for it. Conversely, some of your most responsible actions may be highly predictable by people who know you well, without any examination of your brain. Continuity in the mental processes leading up to a decision is not merely permissible but positively desirable if you are to be regarded as responsible. Ideally, indeed, most of us like to be able to confirm our responsible decisions by going over the ground more than once, to see whether we come to the same conclusion.⁶ If, however, before you made up your mind, there existed a determinate and unique specification of the outcome which could command the assent of everyone, including yourself, whether you knew it or not or liked it or not, then the case would be different. Here we would have to say that the outcome was not only predictable by others, but also *inevitable by you*, and you could not be held to have acted 'freely'.

To call an action 'free' in this sense is therefore to deny the existence of any determinate specification that is binding on (valid and definitive for) everyone, *including the agent*, before he makes up his mind. It is this kind of freedom that I suggest underlies human responsibility. I hope to show that the argument we have been following not only establishes the freedom of normal human choices in this sense, but also offers a mechanistic criterion of diminished responsibility in abnormal conditions, especially in mental disease.

7. What Mechanistic Brain Theory Does Not (and Cannot) Imply

Consider the two situations (i) before 7 p.m. and (ii) at 7 p.m., and the

possibilities (a) and (b) discussed earlier. First, regardless of any brain theory, if a prediction specifies that you will take the decision at (*and not before*) 7 p.m. to go for your train, then clearly you must be wrong to believe it as *certain and inevitable* beforehand, since believing it as certain and inevitable would amount to deciding the outcome *before* 7 p.m.—which would refute the prediction! Before 7 p.m., since we are assuming that what you believe is rigorously represented by the state of your brain, your brain is in a state of type (a), and its future state at 7 p.m. cannot be fully defined in any way that could be believed by you without self-falsification.

At 7 p.m., on the other hand, you are ready to decide. It is no longer self-contradictory for you to accept either *G* or *W*; but you feel (indeed you might claim to know) that both are *open* to you. The crucial question is whether our mechanistic brain theory implies that you are mistaken. As we might by now expect, it does not. Your situation at 7 p.m. fits precisely the case we labelled (b), where there exists a definitive description that can be believed by you without self-falsification—but it is one that could also be rejected by you without any inconsistency, because it will be correct *only* if you accept it. It has no 'take it or leave it' inevitability for you, even on mechanistic assumptions. All that our mechanistic theory can do is to confirm your subjective feeling that your choice of *G* or *W* is genuinely undetermined: an event which depends on your cognitive attitude, as distinct from the myriad other bodily happenings (for example in your digestive system) which are unaffected

by your attitude to them and in that sense outwit your control.

Case (b), in fact, can be generalized. With a given initial state, on the same mechanistic assumptions, there may exist not just one but a whole range of potential state-descriptions, any one of which would become correct *if accepted by the subject*. From a mechanistic point of view, this will happen whenever, given this acceptance, the equation of state has several self-consistent solutions. From the standpoint of the subject, it means that a corresponding range of options are all really *open* to him, so that there is no one outcome which he would be correct to regard as the only possible one. It would be equally consistent with mechanistic theory to accept any one of them as describing his immediate action, and whichever he accepts will in fact determine the action.

So, coming back to our imaginary Determinist and his ciné film, we reach the following conclusions: (1) His evidence proves that *he* was correct, and indeed bound by mechanistic considerations, to expect you to choose G, but (2) it does not retrospectively prove that *you* were mistaken beforehand in believing your action to be undetermined, because (3) though his prediction is demonstrably binding on him, you would have been mistaken in believing it before you made up your mind. Moreover, (4) when you came to make up your mind at 7 p.m., the different possibilities you contemplated were still demonstrably open to you on his mechanistic assumptions; so that (5) even if you were disposed to deny that you were free to choose between them, his mechanistic evidence could in principle demonstrate the reality and extent of your freedom.

8. Diminished Responsibility

As I have already hinted, the last point (5) has a double implication. The logical indeterminacy of a choice for the chooser is demonstrable on the assumption that the process determining its outcome is causally linked, in the embarrassing way we have discussed, to his cognitive process. It follows conversely, however, that where disease, brain damage, *force majeure* or the like override, reduce or abolish the normal causal coupling between the cognitive mechanism and the choosing process, the indeterminacy of the choice for the agent is correspondingly reduced or abolished, and so is his freedom to have it otherwise.

The amount of this reduction seems to offer a possible (and much-needed) objective criterion of *diminished responsibility*.⁷ In court especially, the impression seems widespread today that if a mechanistic explanation can be found for anyone's action, he is automatically to be excused responsibility for it. If a criminal, he must be judged 'sick, not guilty'.

The fallacy should now be apparent. Even when a detached observer can predict an action with complete certainty, this does nothing of itself to prove that the *agent* had no freedom to do otherwise. The key question to be asked of the forensic brain scientist is not whether the action had a physical cause, but whether it admitted of a 'take it or leave it' prediction, which was valid for (binding upon) the agent, whether he knew it or not or liked it or not. I am not suggesting that this would always be an easy question to answer; but at least (if I am right) it directs our attention towards the relevant evidence.

What we should seek to discover from our psychiatric examination is how far the causal connection between the agent's cognitive mechanism and the action in question was deficient at the time in question. The weaker the coupling, the more binding the prediction on the agent, and the less his responsibility.

Of course, our present brain science does not come within miles of the accuracy necessary for such a task; but since it is already being invoked, crude as it is, we need all the more urgently to clarify the principles on which it could be asked to estimate criminal or any other responsibility. To diagnose a prisoner as neurotic, or to discover gross defects in his upbringing, may be highly relevant to the choice of an appropriate *sentence*; but it does nothing of itself to justify denying him his responsibility as a human being, with a right to be held to account for his actions.*

9. Logical Relativity

After such a succession of surprising conclusions, it is high time for some logical stocktaking. We grant that the Determinist with the ciné film was right in believing his prediction before you made up your mind; yet you would have been wrong to believe it, and deterministic brain science had nothing final to tell you—no prediction waited undiscovered—that could rightly have taken the place of your belief that you were taking a free and undetermined choice. Surely *something* is wrong here?

I have, of course, a vested interest

* In the foregoing section I have made use of some passages from an article contributed to a series on 'The Human Brain' in *Science Journal*, May 1967, shortly to be published by Iliffe in book form.

in denying that there is any flaw in the logic of our argument; but I think it is possible to find something wrong in our underlying habits of thought which may fully explain our mental discomfort. This is the common presupposition, normally so taken for granted as to pass unnoticed, that if *anyone* is right in believing something then *everyone* would be right in believing it. A major test of whether a man believes rightly, we feel, is whether other people would be right to share his belief. This is closely related to the idea that whatever is rightly believed (by anyone) can always be expressed in *propositions* which are 'true for all if true for one'. The very idea that two people may have to *differ* in what they believe, in order that both may believe rightly, seems self-contradictory.

But once dragged into the light of day, is this presupposition at all defensible as an absolute rule? To be sure it works well and self-consistently over a wide domain of public discourse *excluding human agents*. Physical relativity theory has cast a shadow over it in one restricted field, but common sense may feel this as too remote to be radically upsetting. Apparent exceptions, such as the nursery argument 'I'm John and you're Mary'—'No, *you're* John and *I'm* Mary', are easily disposed of by saying that each party is making the same *assertion*, though expressing it in appropriately different conventional terms. Even a statement like 'John is mistaken in thinking *P*', although John himself could not believe it without making it out of date, can be dealt with by translating it as '*P* is false, and at this moment John believes *P*'. John can without self-contradiction accept what is *asserted* here, though

he can *express* it only in the past tense. Finally, we might argue, it would surely destroy all rational discussion and objective checking if what John is right to believe is not always identical with what Mary or anyone else is right to believe.

Let us start with the last point. Admittedly it sounds well, and the alternative might seem to be chaos. But of course it is not so. Suppose that Mary's belief is about John's future action. Then, as we have seen, part of the evidence includes the proviso that John will not know or believe the evidence before the time in question. Before the event, comparing notes for agreement is not only impracticable but self-vitiating; and after the event, although John may accept all of Mary's data, he will be unable to derive from them a definitive belief which he would have been correct to hold before the event.*

Chaos? Not a bit of it. On the contrary, from Mary's view of the situation it should be possible for us to infer rigorously what John would be correct to believe. This belief that is correct for John must differ radically from Mary's, in what it affirms as well as in the way it is expressed; but the crucial point is that corresponding to the belief which is valid-for-Mary only *one* belief is valid-for-John and *vice versa*.† In that sense, the two are *complementary correlates*.‡ Though necessarily different, they are not contradic-

tory but *mutually confirmatory as believed by John and Mary respectively*. The obvious name for this principle, that two individuals may have to believe differently in order that each may believe rightly, would seem to be the Principle of Logical Relativity;³ but we must guard against any suggestion of arbitrariness that this name might convey. As we have seen, John is not free to believe *anything he likes*, no matter what Mary believes. Moreover, objectivity is reaffirmed when we realize that John and Mary can agree (in retrospect) that each was right in believing what he or she did, and can appeal to the same recorded data in order to reach this common mind.

Notice, however, the change of emphasis. If we want to retain objectivity (as surely we must) then what has to be settled between John and Mary is not the objective truth of propositions expressing their respective beliefs, but the objectively demonstrable correctness of John's-believing and Mary's-believing. This, after all, is surely the real object of every argument: to decide whether or not *people* are justified in believing what they do.

That we can often achieve this by debating the abstract truth of propositions need not surprise us: for the vast bulk of typical logical argument concerns matters safely remote from the danger area of the participant's own cognitive process. But our normal success must not allow us to forget the restrictions that make it possible, nor blind us to what we are really trying to do. We tend to think of the search for truth as by definition the attempt to discover true propositions, binding on all. It may require considerable and prolonged mental effort to realize that in connection with ourselves this

* Notice here the contrast with the case of 'John is mistaken in thinking *P*'.

† This is not meant to preclude the possibility that the same mental state might be represented by different physical configurations at an atomic or cellular level; but at the level used for Mary's mechanistic explanation these would be equivalent.

definition is inapplicable, its goal illusory. About ourselves the whole truth is not expressible in any single definitive proposition (however lengthy) that all would be correct to believe. If we feel committed in honesty to search for the ultimate truth about ourselves as well as about the world around us, we must be clear that our goal is not some vast *description* at present hidden from us. Whatever else 'ultimate truth' means, it cannot mean that.

It should perhaps be emphasized, however, that there is no question here of abandoning traditional logical reasoning. On the contrary, it is standard propositional logic that we have used to establish the conclusion that not all valid believing consists in the acceptance of universally valid propositions. We are advocating no substitute for propositional logic, but only a recognition that propositions must in general be *examined for their standpoint* if the logical process is to be completed self-consistently, and that universal agreement is possible in principle only about the rightness or wrongness of *people's believing*. To the question *what John would be right to believe* about his future action there can be a single answer which is true no matter who believes it. But to the question *what form John's future action will take* there is no single answer which is true no matter who believes it.

10. Brain-Talk and Mind-Talk

This suggests that we had better take a closer look at the relation we have assumed between the observer's view and the agent's view of personal action. We have spoken of our mythi-

cal Determinist producing a ciné film which proved that he could predict your decision before you made it. Strictly, however, we should have been more careful. What he predicted was, of course, the state of *your brain*. All his data were observations of the physical state of your nervous system and its physical environment.

Now we have granted for the sake of argument the assumption that what a man believes, sees, feels or thinks is rigorously represented by the state of his brain, so that someone who 'knew the code' could 'read off' what goes on in his mind from the state of his brain. (In practice, of course, this is ludicrously beyond the bounds of possibility, but that is beside the present point.) We have thus granted our Determinist the power to predict your *account* of your decision and its motives. Corresponding to your personal account (if truthful) he would have a corroborative mechanical account of your brain-processes.

The crucial point, however, is that although these two stories, one in 'brain-talk' and the other in 'mind-talk', are *correlates* of one another, they are not *translations* of one another, any more than a physicist's description of the electrical currents in a telephone line is a translation of what the speaker is saying. The speaker for example may be complaining that the weather is wet. This statement can be translated into dozens of other languages and sound quite different in each; but in all of them *what is asserted* will be the same; namely, something *about the weather*. The physicist, on the other hand, may also use English or a dozen other languages for his description; in all of them what is asserted will be the same; but

it will be *about the currents in the line*, not about the weather, even though someone who knows the code may read off the one from the other.

Similarly when the Determinist produces his story about your brain, and you produce yours about your decision, it would be a philosophical blunder to imagine that his story could be a *translation* of yours. Indeed, as we have seen, this is so far from being the case that it is possible for his story to be deterministic where yours is indeterministic, without mutual contradiction (indeed with mutual corroboration). Although both relate in a sense to the same events, they are not making assertions about the same aspects of them.

The importance of this emerges when we ask in what terms you and the Determinist can reach agreement in retrospect. If we were to slip into treating his story as a translation of yours, it would seem impossible for him to grant that you were free without denying (by 'translation') that he could predict the state of your brain. Once we realize that the two stories are not translations but correlates, this dilemma can be seen as the *non sequitur* it is. It would equally make no sense for you to claim, conversely, that your *brain* is 'free' because you are. It is not brains, but people, who are (or are not) free. In all discussions of this sort, we must scrupulously distinguish the concepts of brain-machinery from those appropriate to conscious experience, and avoid careless mixing of the two.⁹

11. Predestination

Finally, to tiptoe in where Eddington feared to tread, let us ask how our general thesis would fare if we were

to substitute for physical determinism the theological doctrine of Divine Predestination. Frequently attributed to Calvin and the Presbyterians, this doctrine is in fact explicit in many passages of the Old and New Testaments, particularly the Epistles of Peter and Paul,^{10, 11} where God, as Creator of our whole space-time order, is declared to be sovereign over every event in it, as a potter over his clay, or (to use Dorothy Sayers's well-known simile) as a human novelist over the history he conjures into being.¹²

Surely, it may be said, God's knowledge of us provides an absolute frame of reference which abolishes the logical relativity we considered earlier? True, God is not declared to be a quasi-physical agent *among* other agencies of the created world, but rather the still more sovereign Conceiver of our whole past, present and future history, 'in Whom we (with the whole created order) live and move and have our being'. Yet surely then before we make up our minds, there exists a true description of our decision which God already knows, and which proves our sense of freedom to be an illusion based on ignorance of the truth?

It may help our thinking if we look first at the logical relation between the 'predestined course' of a human novel, and what the people in the story would have been correct to believe. The human author is sovereign over the whole of their created history, which for him is one coherent spatio-temporal fact. No event appears in it that he has not ordained, and we can share his complete knowledge of every decision made by his creatures. This does not mean that we know *in advance* an event still in the future for

us, nor are the events over which the author is sovereign ahead of *him* in (his) time; so for us and him the prefix 'fore-' in such terms as 'fore-known' and 'foreordained' would be inappropriate. All events in the synthetic history, from beginning to end, are for us not prospects but data.

Suppose now that we ask what the people in the story would be correct to believe about the events in their future. John, let us suppose, makes up his mind to ask Mary rather than Jane to marry him. Mary eventually decides to turn him down. Can we say that before either of these decisions was taken, the outcomes were already foreordained and therefore inevitable for John and Mary? Can we say that Mary was not really free to make up her mind, and that she was forced by her author to reject John?

Fortified by our earlier discussions, we can easily see that no such conclusions need follow. No doubt the author *could* write a story in which the characters were mindless puppets, but this is not our story. If John and Mary have been conceived by their author as normal human people, then their decisions will depend upon their cognitive processes in just the way we have already analysed, with the same logical consequences. Not even the author himself can produce a prediction of the outcome which John and Mary would be correct to accept as inevitable before they make up their minds; for no such prediction exists. On such matters our 'predestinarian' knowledge generates nothing definitive for them to know until they make up their minds—and then of course they know the outcome without our help!

On the other hand, if we ask whether John and Mary could be cor-

rect *in retrospect* to acknowledge the sovereignty of their author over their history, the answer seems to be yes. They are no more logically precluded from doing so than you were from accepting the Determinist's ciné film in retrospect. It may be impossible for them to acquire this knowledge; but the point is that in retrospect it would be valid (though mysterious) knowledge for them if they did have it. Moreover, they could most naturally express it by the use of the prefix 'fore-' and speak of their author's having '*foreknown*' and '*foreordained*' what nevertheless they themselves freely determined.

So their author's creative sovereignty and 'fore-knowledge', complete though it is, does not imply that while they are making up their minds there exists a secret prediction of their decision, already known to their author, which they would be correct to believe as certain whether they liked it or not, and which thus proves the outcome to be inevitable for them; nor are they *forced* by their author to make the decisions they do, nor is their responsibility for them in any way reduced. But it does imply that although their author may be unknowable by them, they owe their whole existence and that of their world to him, since he decrees every fact of their history—including the fact that they themselves are free agents.

12. God as Agent

In so far as the parallel holds between human and Divine authorship, the foregoing analysis may help to draw the fatalistic sting commonly felt in the theological doctrine of predestination; but we may still feel that if God is our sovereign Author, we

cannot be held really responsible—at least not by Him. Moreover, as far as our illustration carries us, it might suggest that God must be as unknowable for us as a human novelist is for his characters.

This is not the occasion for a treatise on the deepest mystery of Christian theology, even if I were capable of it; but our story illustration is imperfect at a key point. The significant thing about the Author of our existence, according to Biblical theism, is that He has made Himself also an agent *within* His creation. He addresses men, lays claims upon them, offers them forgiveness and friendship, holds them responsible for their attitude towards Him and what He has revealed of His will. In short, He is declared to be a *participant in dialogue* with us, as well as our Creator.

This, of course, makes all the difference. Dialogue, as we have seen, is a relationship in which the parties necessarily know one another as undetermined agents. If then our Creator has chosen to offer Himself to us in that relationship, *even He must know us as free and responsible beings*. So in any discussion of the relationship between ourselves and God, we must sharply distinguish between what may be logically applicable to 'God-in-dialogue' and to 'God-in-eternity' respectively, just as we would have to do if we were speaking of a human author who had created a history in which he himself was one of the characters. This distinction has some suggestive resemblances to that which Christian theology maintains between the different 'Persons' of the Godhead; but for the present let me only re-emphasize the fact out of which it arises. Logically, we can depend for our existence upon the 'creative' will

of God-in-eternity, and still be answerable for our response as free beings to the 'normative' will of God-in-dialogue. Whatever our attitude to the twin doctrines of divine sovereignty and human responsibility, no logical antithesis can be sustained between them in their Biblical form.

13. Freedom to Choose vs. Freedom to Change

Perhaps at this stage in our long journey someone may wonder whether we have not proved too much. The normal choices of normal human agents, we have shown, are free; but surely from time to time we all know too well the feeling expressed by the Apostle Paul:

'The good that I would I do not;
but the evil which I would not, that I do.'

Are we ever *really* free?

What we have in mind in asking this, I think, is not so much the freedom to opt for what we want, as the freedom to want what we ought, at the right time. It is a question of freedom to *change* rather than freedom to choose; of freedom to become different from what we are, and perhaps more like what in our better moments we want to be.^{6, 13}

It is indeed true to both Christian teaching and common experience that in this latter sense our freedom is severely limited; but this conflicts with nothing in our earlier analysis. On the contrary, if we apply our standard criterion and ask how closely a man's cognitive process is coupled with his various fundamental attitudes and desires, the answer in each case will indicate how far he can be considered

free to change them.* The 'coupling' here will obviously depend on his detailed knowledge of how to go about his purpose. In the concert hall, a sneeze which would otherwise be inevitable may cease to be so if you remember that it can be inhibited by pressing upwards under the nose; if you are free to do this, you are free not to sneeze. If you have almost succeeded in breaking the smoking habit, but suspect that in the company of smokers you will find the urge reinforced, your freedom to abstain may in part be determined by your freedom to choose your company.

In a theological context the same principle applies. The Bible (not just John Calvin) explicitly asserts that by nature we have a bias to sin and are not free to make ourselves good rather than evil; but it goes on to declare that because means are freely available (God offers His own power) to make us different, we can reasonably be held responsible for lingering in our natural condition. Here once again our criterion of freedom accords with what theology and Christian experience mutually affirm.¹³

14. Mechanism in Perspective

It will, I hope, be abundantly clear that in this whole discussion I have had no intention of urging a *belief* in mechanistic determinism, or even in computer models of the brain. My aim has rather been to loosen up a

knotted skein of improperly linked questions in order that each may be disentangled and considered on its merits. I suggest in particular that the question whether all human brain activity has a mechanistic explanation is one we can peacefully leave open for future investigation, no matter how high a view we take of man's power of decision and its moral and religious significance. A complete mechanistic explanation of the brain would not eliminate our freedom, and those who urge mechanistic behaviourism so as to abolish moral and spiritual categories seem to be pursuing an illusion.

Since we started from mechanistic considerations, we have talked much of cognitive mechanisms, and the logical indeterminacy of their states for the agent and those in dialogue with him; but the logical argument could equally have been sustained at the psychological rather than the physical level, with the same result. It is certainly not necessary to *postulate* a mechanistic brain theory in order to demonstrate human freedom! Our aim has been to show that it is equally unnecessary for that purpose to repudiate it.

People may of course define freedom in different ways, and it is open to anyone to say that an action is not 'free', in *his* sense, if even one detached observer could predict it. To him I think we should say that if all he wants is some prediction which the observer in isolation would be right to believe but the agent would be mistaken to accept as inevitable, and which is not binding even on the observer if dialogue is possible with the agent—then he is welcome to it. Such a prediction has no logical cash-value in any context in which responsibility is assessed. We might of course add

* In other words, our freedom here is limited by the options that are *open* to us, not by our power to choose between those that *are* open. It is not the physical determinateness of our brains, but their detailed organization, that frustrates our desires.

that in real life thermodynamic and quantum fluctuations are likely to rob him of his prize even before he sets up his measuring equipment, the data he needs are likely to be perverted by the disturbance of the brain required to measure them, and the complexity of the calculations may well exceed the computing capacity of the universe; but if he is at all a realist he may have capitulated before that point.

To sum up, I am suggesting that fears of mechanistic explanations of brain function are groundless, not because we can be sure that the brain is not a machine, but because even if it were, the whole constellation of claims regarding our inner nature and significance and destiny expressed in our moral tradition and the Christian religion would remain unaffected. It is not people, but brains, that may, or may not, be machines. It is not brains, but people, that choose, freely or otherwise, and in so doing determine their eternal destiny.

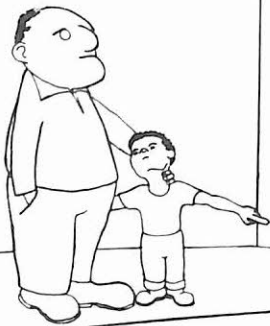
References and Notes

1. EDDINGTON, A. S., *The Nature of the Physical World*, pp. 282–84. Everyman Edition.
2. EDDINGTON, A. S., *The Philosophy of Physical Science*, pp. 183–84. Cambridge University Press.
3. MACKAY, D. M., On the Logical Indeterminacy of a Free Choice. *Proc. XIIIth Int. Congress of Philosophy, Venice*, III (1958), 249–56. Also *Mind*, xix (1960), 31–40.
4. MACKAY, D. M., Man as Observer-Predictor. *Man in his Relationships* (H. Westmann, ed.), pp. 15–28. Routledge, London, 1955.
- Communication and Meaning—a functional approach. *Cross-Cultural Understanding: Epistemology in Anthropology* (F. S. C. Northrop and Helen Livingston, eds.), pp. 162–79. Harper and Row, 1964.
- Information and Prediction in Human Sciences. *Information and Prediction in Science* (S. Dockx and P. Bernays, eds.), Symposium of Int. Acad. for Phil. of Science, 1962. Academic Press, New York, 1965.
5. POPPER, K. R., Indeterminism in Quantum Physics and in Classical Physics. *Brit. J. Phil. Sci.* (1950), I, Part I, pp. 117–33, Part II, pp. 173–95.
6. For a penetrating discussion of the analogy in this respect between deciding and learning, see Vickers, G., *The Art of Judgment—A Study of policy making*. Chapman and Hall, 1965.
7. MACKAY, D. M., Information Theory and Human Information Systems. *Impact of Science on Society*, VIII (1957), 86–101. Also in *Readings in Psychology*, pp. 214–35 (J. Cohen, ed.). Allen and Unwin, 1964.
8. MACKAY, D. M., Complementarity II. *Aristotelian Soc. Suppl.* xxxii (1958), pp. 105–22. The complementarity of causal and subjective accounts of volition was first pointed out by Niels Böhr. See *Dialectica*, II (1948), 318.
9. MACKAY, D. M., The Use of Behavioural Language to Refer to Mechanical Processes. *Brit. J. Phil. of Sci.*, xiii (1962), 89–103.
- A Mind's Eye View of the Brain. *Cybernetics of the Nervous System* (N. Wiener and J. P. Schädé, eds.), *Progress in Brain Research*, xvii (1965), 321–32. Elsevier.

10. I Peter i. 2, 20 and *passim*.
11. Romans viii. 29-30 and ix *passim*.
Ephesians i. 4, 5.
12. SAYERS, DOROTHY, *The Mind of the Maker*. Methuen, 1941.
13. Romans vii. 15-25.
14. Some future states may of course be only trivially dependent on the present state. In general the effect referred to will increase as the future state approaches.

PLEASE DO NOT
TEASE THE ANIMALS

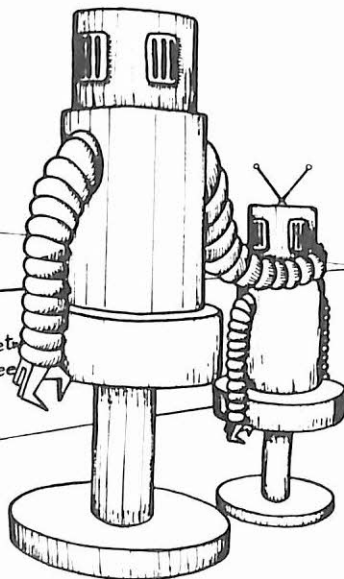
... And these beasts at one stage of our development even surpassed *us* intellectually, but the remarkable fact is that the basis of their intellect is not in a magnetic core, but in *Protoplasm!!*



Homo sapiens
Omnivorous, tribal, evolutionary link between
protoplasm & early computers—see
window to the left



Gorilla gorilla gorilla
Vegetarian



PART THREE

the structure of knowledge

This rather large part consists of a collection of articles focusing on the organization of thinking and behavior. Turing's paper raises the passion-provoking question, "Can a computer think?" and then tries to change that question into an empirically answerable one. Tinbergen looks at one particular bit of knowledge—knowledge about how to find one's way home, in the case of one very simple species: the bee-hunter, *philanthus*. The remaining articles are concerned in one way or another with the usefulness of the now traditional stimulus-response connection theory of behavior. For some kinds of behavior, such a telephone switch-board theory seems reasonable. A stimulus, like a tap on the cheek, impinges on an organism. As a result, a response, like an eye-blink, oc-

curs. Could it be argued that all behavior is made up of simple connections between elementary stimuli and elementary responses? Before reading the following articles you should take a few minutes to ponder this question. Do you think that much of *your* behavior and past learning can be described as stimulus-response connections? Can *all* of it be so expressed? If not, what kind of model seems more appropriate?

INTRODUCTION

In this exciting article, Turing takes a hard look at the question "Can machines think?" Before you read further in this introduction, or read the article, we urge you to stop for a moment and make some notes on how you would go about answering this question. Put differently, Turing is asking, "What is the essential difference between men and machines, if any?" It has often been assumed that men think and machines don't. Turing tries to analyze what, if anything, might be meant by such a statement.

Suppose we were to bring a child of ten years to you and say, "Here is a fine young lad who is healthy in every respect and quite able to get about in the world. He has only one major drawback: he cannot think. He is able to talk and to listen but we have been able to find absolutely no evidence of any thinking processes. We would like you to examine this child and see if you can confirm or refute this extreme diagnosis." In order to check our ludicrous proposition that this child was unable to think, you would probably carry out an extended dialogue with him asking a variety of questions and listening carefully to his answers. If the answers were suitable (although that label is hard to define), you would surely question our assertion that this child is unthinking and try to convince us of the error of our ways. As you shall see, Turing would like to bring such a child to you but, instead of being a ten-year-old lad, he is a computer. The test which Turing proposes in this article is now commonly referred to as "Turing's Test," and has been a source of continued and rich speculation ever since it was first suggested.

As a student of psychology, you will probably be interested in the distinction that Turing draws between emotional and unemotional learning. This dichotomy corresponds rather closely to the separation in experimental psychology between operant or reinforced learning and cognitive or purely informational learning.

At one point in his article Turing refers to *psychokinesis*. This is the hypothesized ability of certain people to be able to cause distant objects to move in various fashions. For example, a gambler with psychokinetic

powers is able to control the outcome of coin tossing, dice rolling, or roulette games. Psychokinesis is obviously of substantial advantage in such enterprises and a power most of us wish we possessed!

COMPUTING MACHINERY AND INTELLIGENCE

A. M. Turing

1. The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the "imitation game." It is played with three people, a man (A), a woman (B), and an interrogator (C)

who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A." The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

Now suppose X is actually A, then A must answer. It is A's object in the game to try and cause C to make the wrong identification. His answer might therefore be:

"My hair is shingled, and the longest strands are about nine inches long."

In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. Alternatively the questions and answers can be repeated by an intermediary. The object of the game for the third player

(B) is to help the interrogator. The best strategy for her is probably to give truthful answers. She can add such things as "I am the woman, don't listen to him!" to her answers, but it will avail nothing as the man can make similar remarks.

We now ask the question, "What will happen when a machine takes the part of A in this game?" Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, "Can machines think?"

2. Critique of the New Problem

As well as asking, "What is the answer to this new form of the question?" one may ask, "Is this new question a worthy one to investigate?" This latter question we investigate without further ado, thereby cutting short an infinite regress.

The new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man. No engineer or chemist claims to be able to produce a material which is indistinguishable from the human skin. It is possible that at some time this might be done, but even supposing this invention available we should feel there was little point in trying to make a "thinking machine" more human by dressing it up in such artificial flesh. The form in which we have set the problem reflects this fact in the condition which prevents the interrogator from seeing or touching the other competitors, or hearing their voices. Some other advantages of the proposed criterion may be shown up by specimen questions and answers. Thus:

Q: Please write me a sonnet on the subject of the Forth Bridge.

A: Count me out on this one. I never could write poetry.

Q: Add 34957 to 70764.

A: (Pause about 30 seconds and then give as answer 105621.

Q: Do you play chess?

A: Yes.

Q: I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play?

A: (After a pause of 15 seconds) R-R8 mate.

The question and answer method seems to be suitable for introducing almost any one of the fields of human endeavour that we wish to include. We do not wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane. The conditions of our game make these disabilities irrelevant. The "witnesses" can brag, if they consider it advisable, as much as they please about their charms, strength or heroism, but the interrogator cannot demand practical demonstrations.

The game may perhaps be criticised on the ground that the odds are weighted too heavily against the machine. If the man were to try and pretend to be the machine he would clearly make a very poor showing. He would be given away at once by slowness and inaccuracy in arithmetic. May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a ma-

chine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection.

It might be urged that when playing the "imitation game" the best strategy for the machine may possibly be something other than imitation of the behaviour of a man. This may be, but I think it is unlikely that there is any great effect of this kind. In any case there is no intention to investigate here the theory of the game, and it will be assumed that the best strategy is to try to provide answers that would naturally be given by a man.

3. The Machines Concerned in the Game

The question which we put in §1 will not be quite definite until we have specified what we mean by the word "machine." It is natural that we should wish to permit every kind of engineering technique to be used in our machines. We also wish to allow the possibility that an engineer or team of engineers may construct a machine which works, but whose manner of operation cannot be satisfactorily described by its constructors because they have applied a method which is largely experimental. Finally, we wish to exclude from the machines men born in the usual manner. It is difficult to frame the definitions so as to satisfy these three conditions. One might for instance insist that the team of engineers should be all of one sex, but this would not really be satisfactory, for it is probably possible to rear a complete individual from a single cell of the skin (say) of a man. To do so would be a feat of biological technique deserving of the very highest praise, but we would not be inclined to regard it as a case of "constructing

a thinking machine." This prompts us to abandon the requirement that every kind of technique should be permitted. We are the more ready to do so in view of the fact that the present interest in "thinking machines" has been aroused by a particular kind of machine, usually called an "electronic computer" or "digital computer." Following this suggestion we only permit digital computers to take part in our game.

This restriction appears at first sight to be a very drastic one. I shall attempt to show that it is not so in reality. To do this necessitates a short account of the nature and properties of these computers.

It may also be said that this identification of machines with digital computers, like our criterion for "thinking," will only be unsatisfactory if (contrary to my belief), it turns out that digital computers are unable to give a good showing in the game.

There are already a number of digital computers in working order, and it may be asked, "Why not try the experiment straight away? It would be easy to satisfy the conditions of the game. A number of interrogators could be used, and statistics compiled to show how often the right identification was given." The short answer is that we are not asking whether all digital computers would do well in the game nor whether the computers at present available would do well, but whether there are imaginable computers which would do well. But this is only the short answer. We shall see this question in a different light later.

4. Digital Computers

The idea behind digital computers may be explained by saying that these

machines are intended to carry out any operations which could be done by a human computer. The human computer is supposed to be following fixed rules; he has no authority to deviate from them in any detail. We may suppose that these rules are supplied in a book, which is altered whenever he is put on to a new job. He has also an unlimited supply of paper on which he does his calculations. He may also do his multiplications and additions on a "desk machine," but this is not important.

If we use the above explanation as a definition we shall be in danger of circularity of argument. We avoid this by giving an outline of the means by which the desired effect is achieved. A digital computer can usually be regarded as consisting of three parts:

- (i) Store.
- (ii) Executive unit.
- (iii) Control.

The store is a store of information, and corresponds to the human computer's paper, whether this is the paper on which he does his calculations or that on which his book of rules is printed. In so far as the human computer does calculations in his head a part of the store will correspond to his memory.

The executive unit is the part which carries out the various individual operations involved in a calculation. What these individual operations are will vary from machine to machine. Usually fairly lengthy operations can be done such as "Multiply 3540675445 by 7076345687" but in some machines only very simple ones such as "Write down 0" are possible.

We have mentioned that the "book of rules" supplied to the computer is replaced in the machine by a part of

the store. It is then called the "table of instructions." It is the duty of the control to see that these instructions are obeyed correctly and in the right order. The control is so constructed that this necessarily happens.

The information in the store is usually broken up into packets of moderately small size. In one machine, for instance, a packet might consist of ten decimal digits. Numbers are assigned to the parts of the store in which the various packets of information are stored, in some systematic manner. A typical instruction might say—

"Add the number stored in position 6809 to that in 4302 and put the result back into the latter storage position."

Needless to say it would not occur in the machine expressed in English. It would more likely be coded in a form such as 6809430217. Here 17 says which of various possible operations is to be performed on the two numbers. In this case the operation is that described above, *viz.*, "Add the number. . . ." It will be noticed that the instruction takes up 10 digits and so forms one packet of information, very conveniently. The control will normally take the instructions to be obeyed in the order of the positions in which they are stored, but occasionally an instruction such as

"Now obey the instruction stored in position 5606, and continue from there"

may be encountered, or again

"If position 4505 contains 0 obey next the instruction stored in 6707, otherwise continue straight on."

Instructions of these latter types are very important because they make it possible for a sequence of operations to be replaced over and over again until some condition is fulfilled, but

in doing so to obey, not fresh instructions on each repetition, but the same ones over and over again. To take a domestic analogy. Suppose Mother wants Tommy to call at the cobbler's every morning on his way to school to see if her shoes are done, she can ask him afresh every morning. Alternatively she can stick up a notice once and for all in the hall which he will see when he leaves for school and which tells him to call for the shoes, and also to destroy the notice when he comes back if he has the shoes with him.

The reader must accept it as a fact that digital computers can be constructed, and indeed have been constructed, according to the principles we have described, and that they can in fact mimic the actions of a human computer very closely.

The book of rules which we have described our human computer as using is of course a convenient fiction. Actual human computers really remember what they have got to do. If one wants to make a machine mimic the behaviour of the human computer in some complex operation one has to ask him how it is done, and then translate the answer into the form of an instruction table. Constructing instruction tables is usually described as "programming." To "programme a machine to carry out the operation A" means to put the appropriate instruction table into the machine so that it will do A.

An interesting variant on the idea of a digital computer is a "digital computer with a random element." These have instructions involving the throwing of a die or some equivalent electronic process; one such instruction might for instance be, "Throw the die and put the resulting number into store 1000." Sometimes such a

machine is described as having free will (though I would not use this phrase myself). It is not normally possible to determine from observing a machine whether it has a random element, for a similar effect can be produced by such devices as making the choices depend on the digits of the decimal for π .

Most actual digital computers have only a finite store. There is no theoretical difficulty in the idea of a computer with an unlimited store. Of course only a finite part can have been used at any one time. Likewise only a finite amount can have been constructed, but we can imagine more and more being added as required. Such computers have special theoretical interest and will be called infinite capacity computers.

The idea of a digital computer is an old one. Charles Babbage, Lucasian Professor of Mathematics at Cambridge from 1828 to 1839, planned such a machine, called the Analytical Engine, but it was never completed. Although Babbage had all the essential ideas, his machine was not at that time such a very attractive prospect. The speed which would have been available would be definitely faster than a human computer but something like 100 times slower than the Manchester machine, itself one of the slower of the modern machines. The storage was to be purely mechanical, using wheels and cards.

The fact that Babbage's Analytical Engine was to be entirely mechanical will help us to rid ourselves of a superstition. Importance is often attached to the fact that modern digital computers are electrical, and that the nervous system also is electrical. Since Babbage's machine was not electrical, and since all digital computers are in a sense equivalent, we see that this

use of electricity cannot be of theoretical importance. Of course electricity usually comes in where fast signalling is concerned, so that it is not surprising that we find it in both these connections. In the nervous system chemical phenomena are at least as important as electrical. In certain computers the storage system is mainly acoustic. The feature of using electricity is thus seen to be only a very superficial similarity. If we wish to find such similarities we should look rather for mathematical analogies of function.

5. Universality of Digital Computers

The digital computers considered in the last section may be classified amongst the "discrete-state machines." These are the machines which move by sudden jumps or clicks from one quite definite state to another. These states are sufficiently different for the possibility of confusion between them to be ignored. Strictly speaking there are no such machines. Everything really moves continuously. But there are many kinds of machine which can profitably be *thought* of as being discrete-state machines. For instance in considering the switches for a lighting system it is a convenient fiction that each switch must be definitely on or definitely off. There must be intermediate positions, but for most purposes we can forget about them. As an example of a discrete-state machine we might consider a wheel which clicks round through 120° once a second, but may be stopped by a lever which can be operated from outside; in addition a lamp is to light in one of the positions of the wheel. This machine could be described abstractly as follows. The internal state of the

machine (which is described by the position of the wheel) may be q_1 q_2 or q_3 . There is an input signal i_0 or i_1 (position of lever). The internal state at any moment is determined by the last state and input signal according to the table

Last State				
q_1	q_2	q_3		
i_0	q_2	q_3	q_1	
Input				
i_1	q_1	q_2	q_3	

The output signals, the only externally visible indication of the internal state (the light) are described by the table

State	q_1	q_2	q_3
Output	o_0	o_0	o_1

This example is typical of discrete-state machines. They can be described by such tables provided they have only a finite number of possible states.

It will seem that given the initial state of the machine and the input signals it is always possible to predict all future states. This is reminiscent of Laplace's view that from the complete state of the universe at one moment of time, as described by the positions and velocities of all particles, it should be possible to predict all future states. The prediction which we are considering is, however, rather nearer to practicability than that considered by Laplace. The system of the "universe as a whole" is such that quite small errors in the initial conditions can have an overwhelming effect at a later time. The displacement of a single electron by a billionth of a centimetre at one moment might make the difference between a man being killed by an avalanche a year

later, or escaping. It is an essential property of the mechanical systems which we have called "discrete-state machines" that this phenomenon does not occur. Even when we consider the actual physical machines instead of the idealised machines, reasonably accurate knowledge of the state at one moment yields reasonably accurate knowledge any number of steps later.

As we have mentioned, digital computers fall within the class of discrete-state machines. But the number of states of which such a machine is capable is usually enormously large. For instance, the number for the machine now working at Manchester is about $2^{165,000}$, i.e., about $10^{50,000}$. Compare this with our example of the clicking wheel described above, which had three states. It is not difficult to see why the number of states should be so immense. The computer includes a store corresponding to the paper used by a human computer. It must be possible to write into the store any one of the combinations of symbols which might have been written on the paper. For simplicity suppose that only digits from 0 to 9 are used as symbols. Variations in handwriting are ignored. Suppose the computer is allowed 100 sheets of paper each containing 50 lines each with room for 30 digits. Then the number of states is $10^{100 \times 50 \times 30}$, i.e., $10^{150,000}$. This is about the number of states of three Manchester machines put together. The logarithm to the base two of the number of states is usually called the "storage capacity" of the machine. Thus the Manchester machine has a storage capacity of about 165,000 and the wheel machine of our example about 1.6. If two machines are put together their capacities must be added to obtain the capacity of the resultant machine. This leads

to the possibility of statements such as "The Manchester machine contains 64 magnetic tracks each with a capacity of 2560, eight electronic tubes with a capacity of 1280. Miscellaneous storage amounts to about 300 making a total of 174,380."

Given the table corresponding to a discrete-state machine it is possible to predict what it will do. There is no reason why this calculation should not be carried out by means of a digital computer. Provided it could be carried out sufficiently quickly the digital computer could mimic the behavior of any discrete-state machine. The imitation game could then be played with the machine in question (as B) and the mimicking digital computer (as A) and the interrogator would be unable to distinguish them. Of course the digital computer must have an adequate storage capacity as well as working sufficiently fast. Moreover, it must be programmed afresh for each new machine which it is desired to mimic.

This special property of digital computers, that they can mimic any discrete-state machine, is described by saying that they are *universal* machines. The existence of machines with this property has the important consequence that, considerations of speed apart, it is unnecessary to design various new machines to do various computing processes. They can all be done with one digital computer, suitably programmed for each case. It will be seen that as a consequence of this all digital computers are in a sense equivalent.

We may now consider again the point raised at the end of §3. It was suggested tentatively that the question, "Can machines think?" should be replaced by "Are there imaginable digital computers which would do

well in the imitation game?" If we wish we can make this superficially more general and ask "Are there discrete-state machines which would do well?" But in view of the universality property we see that either of these questions is equivalent to this, "Let us fix our attention on one particular digital computer *C*. Is it true that by modifying this computer to have an adequate storage, suitably increasing its speed of action, and providing it with an appropriate programme, *C* can be made to play satisfactorily the part of *A* in the imitation game, the part of *B* being taken by a man?"

6. Contrary Views on the Main Question

We may now consider the ground to have been cleared and we are ready to proceed to the debate on our question, "Can machines think?" and the variant of it quoted at the end of the last section. We cannot altogether abandon the original form of the problem, for opinions will differ as to the appropriateness of the substitution and we must at least listen to what has to be said in this connexion.

It will simplify matters for the reader if I explain first my own beliefs in the matter. Consider first the more accurate form of the question. I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words

and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe further that no useful purpose is served by concealing these beliefs. The popular view that scientists proceed inexorably from well-established fact to well-established fact, never being influenced by any improved conjecture, is quite mistaken. Provided it is made clear which are proved facts and which are conjectures, no harm can result. Conjectures are of great importance since they suggest useful lines of research.

I now proceed to consider opinions opposed to my own.

(1) THE THEOLOGICAL OBJECTION

Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or to machines. Hence no animal or machine can think.¹

I am unable to accept any part of this, but will attempt to reply in theological terms. I should find the argument more convincing if animals were classed with men, for there is a greater difference, to my mind, between the typical animate and the inanimate than there is between man and the other animals. The arbitrary character of the orthodox view becomes clearer if we consider how it might appear to a member of some other religious community. How do Christians regard the Moslem view that women have no souls? But let us

¹ Possibly this view is heretical. St. Thomas Aquinas [*Summa Theologica*, quoted by Bertrand Russell (1945, p. 458)] states that God cannot make a man to have no soul. But this may not be a real restriction on His powers, but only a result of the fact that men's souls are immortal, and therefore indestructible.

leave this point aside and return to the main argument. It appears to me that the argument quoted above implies a serious restriction of the omnipotence of the Almighty. It is admitted that there are certain things that He cannot do such as making one equal to two, but should we not believe that He has freedom to confer a soul on an elephant if He sees fit? We might expect that He would only exercise this power in conjunction with a mutation which provided the elephant with an appropriately improved brain to minister to the needs of this soul. An argument of exactly similar form may be made for the case of machines. It may seem different because it is more difficult to "swallow." But this really only means that we think it would be less likely that He would consider the circumstances suitable for conferring a soul. The circumstances in question are discussed in the rest of this paper. In attempting to construct such machines we should not be irreverently usurping His power of creating souls, any more than we are in the procreation of children: rather we are, in either case, instruments of His will providing mansions for the souls that He creates.

However, this is mere speculation. I am not very impressed with theological arguments whatever they may be used to support. Such arguments have often been found unsatisfactory in the past. In the time of Galileo it was argued that the texts, "And the sun stood still . . . and hasted not to go down about a whole day" (Joshua x. 13) and "He laid the foundations of the earth, that it should not move at any time" (Psalm cv. 5) were an adequate refutation of the Copernican theory. With our present knowledge such an argument appears futile. When that knowledge was not avail-

able it made a quite different impression.

(2) THE "HEAD IN THE SAND" OBJECTION

"The consequences of machines thinking would be too dreadful. Let us hope and believe that they cannot do so."

This argument is seldom expressed quite so openly as in the form above. But it affects most of us who think about it at all. We like to believe that Man is in some subtle way superior to the rest of creation. It is best if he can be shown to be *necessarily* superior, for then there is no danger of him losing his commanding position. The popularity of the theological argument is clearly connected with this feeling. It is likely to be quite strong in intellectual people, since they value the power of thinking more highly than others, and are more inclined to base their belief in the superiority of Man on this power.

I do not think that this argument is sufficiently substantial to require refutation. Consolation would be more appropriate: perhaps this should be sought in the transmigration of souls.

(3) THE MATHEMATICAL OBJECTION

There are a number of results of mathematical logic which can be used to show that there are limitations to the powers of discrete-state machines. The best known of these results is known as Gödel's theorem (1931) and shows that in any sufficiently powerful logical system statements can be formulated which can neither be proved nor disproved within the system, unless possibly the system itself is inconsistent. There are other, in some respects similar, results due to

Church (1936), Kleene (1935), Rosser, and Turing (1937). The latter result is the most convenient to consider, since it refers directly to machines, whereas the others can only be used in a comparatively indirect argument: for instance if Gödel's theorem is to be used we need in addition to have some means of describing logical systems in terms of machines, and machines in terms of logical systems. The result in question refers to a type of machine which is essentially a digital computer with an infinite capacity. It states that there are certain things that such a machine cannot do. If it is rigged up to give answers to questions as in the imitation game, there will be some questions to which it will either give a wrong answer, or fail to give an answer at all however much time is allowed for a reply. There may, of course, be many such questions, and questions which cannot be answered by one machine may be satisfactorily answered by another. We are of course supposing for the present that the questions are of the kind to which an answer "Yes" or "No" is appropriate, rather than questions such as "What do you think of Picasso?" The questions that we know the machines must fail on are of this type, "Consider the machine specified as follows. . . . Will this machine ever answer 'Yes' to any question?" The dots are to be replaced by a description of some machine in a standard form, which could be something like that used in §5. When the machine described bears a certain comparatively simple relation to the machine which is under interrogation, it can be shown that the answer is either wrong or not forthcoming. This is the mathematical result: it is argued that it proves a disability of machines to which the human intellect is not subject.

The short answer to this argument is that although it is established that there are limitations to the powers of any particular machine, it has only been stated, without any sort of proof, that no such limitations apply to the human intellect. But I do not think this view can be dismissed quite so lightly. Whenever one of these machines is asked the appropriate critical question, and gives a definite answer, we know that this answer must be wrong, and this gives us a certain feeling of superiority. Is this feeling illusory? It is no doubt quite genuine, but I do not think too much importance should be attached to it. We too often give wrong answers to questions ourselves to be justified in being very pleased at such evidence of fallibility on the part of the machines. Further, our superiority can only be felt on such an occasion in relation to the one machine over which we have scored our petty triumph. There would be no question of triumphing simultaneously over *all* machines. In short, then, there might be men cleverer than any given machine, but then again there might be other machines cleverer again, and so on.

Those who hold to the mathematical argument would, I think, mostly be willing to accept the imitation game as a basis for discussion. Those who believe in the two previous objections would probably not be interested in any criteria.

(4) THE ARGUMENT FROM CONSCIOUSNESS

This argument is very well expressed in Professor Jefferson's Lister Oration for 1949, from which I quote. "Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not

by the chance fall of symbols, could we agree that machine equals brain—that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants.”

This argument appears to be a denial of the validity of our test. According to the most extreme form of this view the only way by which one could be sure that a machine thinks is to *be* the machine and to feel oneself thinking. One could then describe these feelings to the world, but of course no one would be justified in taking any notice. Likewise according to this view the only way to know that a *man* thinks is to be that particular man. It is in fact the solipsist point of view. It may be the most logical view to hold but it makes communication of ideas difficult. A is liable to believe “A thinks but B does not” whilst B believes “B thinks but A does not.” Instead of arguing continually over this point it is usual to have the polite convention that everyone thinks.

I am sure that Professor Jefferson does not wish to adopt the extreme and solipsist point of view. Probably he would be quite willing to accept the imitation game as a test. The game (with the player B omitted) is frequently used in practice under the name of *viva voce* to discover whether some one really understands something or has “learnt it parrot fashion.” Let us listen in to a part of such a *viva voce*:

Interrogator: In the first line of your sonnet which reads “Shall

I compare thee to a summer’s day,” would not “a spring day” do as well or better?

Witness: It wouldn’t scan.

Interrogator: How about “a winter’s day.” That would scan all right.

Witness: Yes, but nobody wants to be compared to a winter’s day.

Interrogator: Would you say Mr. Pickwick reminded you of Christmas?

Witness: In a way.

Interrogator: Yet Christmas is a winter’s day, and I do not think Mr. Pickwick would mind the comparison.

Witness: I don’t think you’re serious. By a winter’s day one means a typical winter’s day, rather than a special one like Christmas.

And so on. What would Professor Jefferson say if the sonnet-writing machine was able to answer like this in the *viva voce*? I do not know whether he would regard the machine as “merely artificially signalling” these answers, but if the answers were as satisfactory and sustained as in the above passage I do not think he would describe it as “an easy contrivance.” This phrase is, I think, intended to cover such devices as the inclusion in the machine of a record of someone reading a sonnet, with appropriate switching to turn it on from time to time.

In short then, I think that most of those who support the argument from consciousness could be persuaded to abandon it rather than be forced into

the solipsist position. They will then probably be willing to accept our test.

I do not wish to give the impression that I think there is no mystery about consciousness. There is, for instance, something of a paradox connected with any attempt to localise it. But I do not think these mysteries necessarily need to be solved before we can answer the question with which we are concerned in this paper.

(5) ARGUMENTS FROM VARIOUS DISABILITIES

These arguments take the form, "I grant you that you can make machines do all the things you have mentioned but you will never be able to make one to do X." Numerous features X are suggested in this connexion. I offer a selection:

Be kind, resourceful, beautiful, friendly, have initiative, have a sense of humour, tell right from wrong, make mistakes, fall in love, enjoy strawberries and cream, make some one fall in love with it, learn from experience, use words properly, be the subject of its own thought, have as much diversity of behaviour as a man, do something really new.

No support is usually offered for these statements. I believe they are mostly founded on the principle of scientific induction. A man has seen thousands of machines in his lifetime. From what he sees of them he draws a number of general conclusions. They are ugly, each is designed for a very limited purpose, when required for a minutely different purpose they are useless, the variety of behaviour of any one of them is very small, etc., etc. Naturally he concludes that these are necessary properties of machines in general. Many of these limitations are associated with the very small

storage capacity of most machines. (I am assuming that the idea of storage capacity is extended in some way to cover machines other than discrete-state machines. The exact definition does not matter as no mathematical accuracy is claimed in the present discussion.) A few years ago, when very little had been heard of digital computers, it was possible to elicit much incredulity concerning them, if one mentioned their properties without describing their construction. That was presumably due to a similar application of the principle of scientific induction. These applications of the principle are of course largely unconscious. When a burnt child fears the fire and shows that he fears it by avoiding it, I should say that he was applying scientific induction. (I could of course also describe his behaviour in many other ways.) The works and customs of mankind do not seem to be very suitable material to which to apply scientific induction. A very large part of space-time must be investigated, if reliable results are to be obtained. Otherwise we may (as most English children do) decide that everybody speaks English, and that it is silly to learn French.

There are, however, special remarks to be made about many of the disabilities that have been mentioned. The inability to enjoy strawberries and cream may have struck the reader as frivolous. Possibly a machine might be made to enjoy this delicious dish, but any attempt to make one do so would be idiotic. What is important about this disability is that it contributes to some of the other disabilities, *e.g.*, to the difficulty of the same kind of friendliness occurring between man and machine as between white man and white man, or between black man and black man.

The claim that "machines cannot make mistakes" seems a curious one. One is tempted to retort, "Are they any the worse for that?" But let us adopt a more sympathetic attitude, and try to see what is really meant. I think this criticism can be explained in terms of the imitation game. It is claimed that the interrogator could distinguish the machine from the man simply by setting them a number of problems in arithmetic. The machine would be unmasked because of its deadly accuracy. The reply to this is simple. The machine (programmed for playing the game) would not attempt to give the *right* answers to the arithmetic problems. It would deliberately introduce mistakes in a manner calculated to confuse the interrogator. A mechanical fault would probably show itself through an unsuitable decision as to what sort of a mistake to make in the arithmetic. Even this interpretation of the criticism is not sufficiently sympathetic. But we cannot afford the space to go into it much further. It seems to me that this criticism depends on a confusion between two kinds of mistake. We may call them "errors of functioning" and "errors of conclusion." Errors of functioning are due to some mechanical or electrical fault which causes the machine to behave otherwise than it was designed to do. In philosophical discussions one likes to ignore the possibility of such errors; one is therefore discussing "abstract machines." These abstract machines are mathematical fictions rather than physical objects. By definition they are incapable of errors of functioning. In this sense we can truly say that "machines can never make mistakes." Errors of conclusion can only arise when some meaning is attached to the output signals from the machine. The

machine might, for instance, type out mathematical equations, or sentences in English. When a false proposition is typed we say that the machine has committed an error of conclusion. There is clearly no reason at all for saying that a machine cannot make this kind of mistake. It might do nothing but type out repeatedly " $0 = 1$." To take a less perverse example, it might have some method for drawing conclusions by scientific induction. We must expect such a method to lead occasionally to erroneous results.

The claim that a machine cannot be the subject of its own thought can of course only be answered if it can be shown that the machine has *some* thought with *some* subject matter. Nevertheless, "the subject matter of a machine's operations" does seem to mean something, at least to the people who deal with it. If, for instance, the machine was trying to find a solution of the equation $x^2 - 40x - 11 = 0$ one would be tempted to describe this equation as part of the machine's subject matter at that moment. In this sort of sense a machine undoubtedly can be its own subject matter. It may be used to help in making up its own programmes, or to predict the effect of alterations in its own structure. By observing the results of its own behaviour it can modify its own programmes so as to achieve some purpose more effectively. These are possibilities of the near future, rather than Utopian dreams.

The criticism that a machine cannot have much diversity of behaviour is just a way of saying that it cannot have much storage capacity. Until fairly recently a storage capacity of even a thousand digits was very rare. The criticisms that we are considering here are often disguised forms of

the argument from consciousness. Usually if one maintains that a machine *can* do one of these things, and describes the kind of method that the machine could use, one will not make much of an impression. It is thought that the method (whatever it may be, for it must be mechanical) is really rather base. Compare the parentheses in Jefferson's statement.

(6) LADY LOVELACE'S OBJECTION

Our most detailed information of Babbage's Analytical Engine comes from a memoir by Lady Lovelace (1842). In it she states, "The Analytical Engine has no pretensions to *originate* anything. It can do *whatever we know how to order it to perform*" (her italics). This statement is quoted by Hartree (1949) who adds: "This does not imply that it may not be possible to construct electronic equipment to which will 'think for itself,' or in which, in biological terms, one could set up a conditioned reflex, which would serve as a basis for 'learning.' Whether this is possible in principle or not is a stimulating and exciting question, suggested by some of these recent developments. But it did not seem that the machines constructed or projected at the time had this property."

I am in thorough agreement with Hartree over this. It will be noticed that he does not assert that the machines in question had not got the property, but rather that the evidence available to Lady Lovelace did not encourage her to believe that they had it. It is quite possible that the machines in question had in a sense got this property. For suppose that some discrete-state machine has the property. The Analytical Engine was

a universal digital computer, so that, if its storage capacity and speed were adequate, it could by suitable programming be made to mimic the machine in question. Probably this argument did not occur to the Countess or to Babbage. In any case there was no obligation on them to claim all that could be claimed.

This whole question will be considered again under the heading of learning machines.

A variant of Lady Lovelace's objection states that a machine can "never do anything really new." This may be parried for a moment with the saw, "There is nothing new under the sun." Who can be certain that "original work" that he has done was not simply the growth of the seed planted in him by teaching, or the effect of following well-known general principles. A better variant of the objection says that a machine can never "take us by surprise." This statement is a more direct challenge and can be met directly. Machines take me by surprise with great frequency. This is largely because I do not do sufficient calculation to decide what to expect them to do, or rather because, although I do a calculation, I do it in a hurried, slipshod fashion, taking risks. Perhaps I say to myself, "I suppose the voltage here ought to be the same as there: anyway let's assume it is." Naturally I am often wrong, and the result is a surprise for me for by the time the experiment is done these assumptions have been forgotten. These admissions lay me open to lectures on the subject of my vicious ways, but do not throw any doubt on my credibility when I testify to the surprises I experience.

I do not expect this reply to silence my critic. He will probably say that such surprises are due to some cre-

ative mental act on my part, and reflect no credit on the machine. This leads us back to the argument from consciousness, and far from the idea of surprise. It is a line of argument we must consider closed, but it is perhaps worth remarking that the appreciation of something as surprising requires as much of a "creative mental act" whether the surprising event originates from a man, a book, a machine or anything else.

The view that machines cannot give rise to surprises is due, I believe, to a fallacy to which philosophers and mathematicians are particularly subject. This is the assumption that as soon as a fact is presented to a mind all consequences of that fact spring into the mind simultaneously with it. It is a very useful assumption under many circumstances, but one too easily forgets that it is false. A natural consequence of doing so is that one then assumes that there is no virtue in the mere working out of consequences from data and general principles.

(7) ARGUMENT FROM CONTINUITY IN THE NERVOUS SYSTEM

The nervous system is certainly not a discrete-state machine. A small error in the information about the size of a nervous impulse impinging on a neuron, may make a large difference to the size of the outgoing impulse. It may be argued that, this being so, one cannot expect to be able to mimic the behaviour of the nervous system with a discrete-state system.

It is true that a discrete-state machine must be different from a continuous machine. But if we adhere to the conditions of the imitation game, the interrogator will not be able to take any advantage of this difference.

The situation can be made clearer if we consider some other simpler continuous machine. A differential analyser will do very well. (A differential analyser is a certain kind of machine not of the discrete-state type used for some kinds of calculation.) Some of these provide their answers in a typed form, and so are suitable for taking part in the game. It would not be possible for a digital computer to predict exactly what answers the differential analyser would give to a problem, but it would be quite capable of giving the right sort of answer. For instance, if asked to give the value of π (actually about 3.1416) it would be reasonable to choose at random between the values 3.12, 3.13, 3.14, 3.15, 3.16 with the probabilities of 0.05, 0.15, 0.55, 0.19, 0.06 (say). Under these circumstances it would be very difficult for the interrogator to distinguish the differential analyser from the digital computer.

(8) THE ARGUMENT FROM INFORMALITY OF BEHAVIOUR

It is not possible to produce a set of rules purporting to describe what a man should do in every conceivable set of circumstances. One might for instance have a rule that one is to stop when one sees a red traffic light, and to go if one sees a green one, but what if by some fault both appear together? One may perhaps decide that it is safest to stop. But some further difficulty may well arise from this decision later. To attempt to provide rules of conduct to cover every eventuality, even those arising from traffic lights, appears to be impossible. With all this I agree.

From this it is argued that we cannot be machines. I shall try to reproduce the argument, but I fear I shall

hardly do it justice. It seems to run something like this. "If each man had a definite set of rules of conduct by which he regulated his life he would be no better than a machine. But there are no such rules, so men cannot be machines." The undistributed middle is glaring. I do not think the argument is ever put quite like this, but I believe this is the argument used nevertheless. There may however be a certain confusion between "rules of conduct" and "laws of behaviour" to cloud the issue. By "rules of conduct" I mean precepts such as "Stop if you see red lights," on which one can act, and of which one can be conscious. By "laws of behaviour" I mean laws of nature as applied to a man's body such as "if you pinch him he will squeak." If we substitute "laws of behaviour which regulate his life" for "laws of conduct by which he regulates his life" in the argument quoted the undistributed middle is no longer insuperable. For we believe that it is not only true that being regulated by laws of behaviour implies being some sort of machine (though not necessarily a discrete-state machine), but that conversely being such a machine implies being regulated by such laws. However, we cannot so easily convince ourselves of the absence of complete laws of behaviour as of complete rules of conduct. The only way we know of for finding such laws is scientific observation, and we certainly know of no circumstances under which we could say, "We have searched enough. There are no such laws."

We can demonstrate more forcibly that any such statement would be unjustified. For suppose we could be sure of finding such laws if they existed. Then given a discrete-state machine it should certainly be possible to discover by observation sufficient

about it to predict its future behaviour, and this within a reasonable time, say a thousand years. But this does not seem to be the case. I have set up on the Manchester computer a small programme using only 1,000 units of storage, whereby the machine supplied with one sixteen-figure number replies with another within two seconds. I would defy anyone to learn from these replies sufficient about the programme to be able to predict any replies to untried values.

(9) THE ARGUMENT FROM EXTRASENSORY PERCEPTION

I assume that the reader is familiar with the idea of extrasensory perception, and the meaning of the four items of it, *viz.*, telepathy, clairvoyance, precognition and psychokinesis. These disturbing phenomena seem to deny all our usual scientific ideas. How we should like to discredit them! Unfortunately the statistical evidence, at least for telepathy, is overwhelming. It is very difficult to rearrange one's ideas so as to fit these new facts in. Once one has accepted them it does not seem a very big step to believe in ghosts and bogies. The idea that our bodies move simply according to the known laws of physics, together with some others not yet discovered but somewhat similar, would be one of the first to go.

This argument is to my mind quite a strong one. One can say in reply that many scientific theories seem to remain workable in practice, in spite of clashing with ESP; that in fact one can get along very nicely if one forgets about it. This is rather cold comfort, and one fears that thinking is just the kind of phenomenon where ESP may be especially relevant.

A more specific argument based on

ESP might run as follows: "Let us play the imitation game, using as witnesses, a man who is good as a telepathic receiver, and a digital computer. The interrogator can ask such questions as 'What suit does the card in my right hand belong to?' The man by telepathy or clairvoyance gives the right answer 130 times out of 400 cards. The machine can only guess at random, and perhaps gets 104 right, so the interrogator makes the right identification." There is an interesting possibility which opens here. Suppose the digital computer contains a random number generator. Then it will be natural to use this to decide what answer to give. But then the random number generator will be subject to the psychokinetic powers of the interrogator. Perhaps this psychokinesis might cause the machine to guess right more often than would be expected on a probability calculation, so that the interrogator might still be unable to make the right identification. On the other hand, he might be able to guess right without any questioning, by clairvoyance. With ESP anything may happen.

If telepathy is admitted it will be necessary to tighten our test up. The situation could be regarded as analogous to that which would occur if the interrogator were talking to himself and one of the competitors was listening with his ear to the wall. To put the competitors into a "telepathy-proof room" would satisfy all requirements.

7. Learning Machines

The reader will have anticipated that I have no very convincing arguments of a positive nature to support my views. If I had I should not have taken such pains to point out the

fallacies in contrary views. Such evidence as I have I shall now give.

Let us return for a moment to Lady Lovelace's objection, which stated that the machine can only do what we tell it to do. One could say that a man can "inject" an idea into the machine, and that it will respond to a certain extent and then drop into quiescence, like a piano string struck by a hammer. Another simile would be an atomic pile of less than critical size: an injected idea is to correspond to a neutron entering the pile from without. Each such neutron will cause a certain disturbance which eventually dies away. If, however, the size of the pile is sufficiently increased, the disturbance caused by such an incoming neutron will very likely go on and on increasing until the whole pile is destroyed. Is there a corresponding phenomenon for minds, and is there one for machines? There does seem to be one for the human mind. The majority of them seem to be "subcritical," *i.e.*, to correspond in this analogy to piles of subcritical size. An idea presented to such a mind will on average give rise to less than one idea in reply. A smallish proportion are supercritical. An idea presented to such a mind may give rise to a whole "theory" consisting of secondary, tertiary and more remote ideas. Animals' minds seem to be very definitely subcritical. Adhering to this analogy we ask, "Can a machine be made to be supercritical?"

The "skin-of-an-onion" analogy is also helpful. In considering the functions of the mind or the brain we find certain operations which we can explain in purely mechanical terms. This we say does not correspond to the real mind: it is a sort of skin which we must strip off if we are to find the real mind. But then in what

remains we find a further skin to be stripped off, and so on. Proceeding in this way do we ever come to the "real" mind, or do we eventually come to the skin which has nothing in it? In the latter case the whole mind is mechanical. (It would not be a discrete-state machine however. We have discussed this.)

These last two paragraphs do not claim to be convincing arguments. They should rather be described as "recitations tending to produce belief."

The only really satisfactory support that can be given for the view expressed at the beginning of §6, will be that provided by waiting for the end of the century and then doing the experiment described. But what can we say in the meantime? What steps should be taken now if the experiment is to be successful?

As I have explained, the problem is mainly one of programming. Advances in engineering will have to be made too, but it seems unlikely that these will not be adequate for the requirements. Estimates of the storage capacity of the brain vary from 10^{10} to 10^{15} binary digits. I incline to the lower values and believe that only a very small fraction is used for the higher types of thinking. Most of it is probably used for the retention of visual impressions. I should be surprised if more than 10^9 was required for satisfactory playing of the imitation game, at any rate against a blind man. (Note: The capacity of the *Encyclopaedia Britannica*, 11th edition, is 2×10^9 .) A storage capacity of 10^7 would be a very practicable possibility even by present techniques. It is probably not necessary to increase the speed of operations of the machines at all. Parts of modern machines which can be regarded as analogs of

nerve cells work about a thousand times faster than the latter. This should provide a "margin of safety" which could cover losses of speed arising in many ways. Our problem then is to find out how to programme these machines to play the game. At my present rate of working I produce about a thousand digits of programme a day, so that about sixty workers, working steadily through the fifty years might accomplish the job, if nothing went into the wastepaper basket. Some more expeditious method seems desirable.

In the process of trying to imitate an adult human mind we are bound to think a good deal about the process which has brought it to the state that it is in. We may notice three components.

- (a) The initial state of the mind, say at birth,
- (b) The education to which it has been subjected,
- (c) Other experience, not to be described as education, to which it has been subjected.

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed. The amount of work in the education we

can assume, as a first approximation, to be much the same as for the human child.

We have thus divided our problem into two parts. The child programme and the education process. These two remain very closely connected. We cannot expect to find a good child machine at the first attempt. One must experiment with teaching one such machine and see how well it learns. One can then try another and see if it is better or worse. There is an obvious connection between this process and evolution, by the identifications

Structure of the child machine =
hereditary material
Changes of the child machine =
mutations
Natural selection =
judgment of the experimenter

One may hope, however, that this process will be more expeditious than evolution. The survival of the fittest is a slow method for measuring advantages. The experimenter, by the exercise of intelligence, should be able to speed it up. Equally important is the fact that he is not restricted to random mutations. If he can trace a cause for some weakness he can probably think of the kind of mutation which will improve it.

It will not be possible to apply exactly the same teaching process to the machine as to a normal child. It will not, for instance, be provided with legs, so that it could not be asked to go out and fill the coal scuttle. Possibly it might not have eyes. But however well these deficiencies might be overcome by clever engineering, one could not send the creature to school without the other children making excessive fun of it. It must be given home tuition. We need not be too

concerned about the legs, eyes, etc. The example of Miss Helen Keller shows that education can take place provided that communication in both directions between teacher and pupil can take place by some means or other.

We normally associate punishments and rewards with the teaching process. Some simple child machines can be constructed or programmed on this sort of principle. The machine has to be so constructed that events which shortly preceded the occurrence of a punishment signal are unlikely to be repeated, whereas a reward signal increased the probability of repetition of the events which led up to it. These definitions do not presuppose any feelings on the part of the machine. I have done some experiments with one such child machine, and succeeded in teaching it a few things, but the teaching method was too unorthodox for the experiment to be considered really successful.

The use of punishments and rewards can at best be a part of the teaching process. Roughly speaking, if the teacher has no other means of communicating to the pupil, the amount of information which can reach him does not exceed the total number of rewards and punishments applied. By the time a child has learnt to repeat "Casabianca" he would probably feel very sore indeed, if the text could only be discovered by a "Twenty Questions" technique, every "NO" taking the form of a blow. It is necessary therefore to have some other "unemotional" channels of communication. If these are available it is possible to teach a machine by punishments and rewards to obey orders given in some language, *e.g.*, a symbolic language. These orders are to be transmitted through the "unemotional" channels. The use of this

language will diminish greatly the number of punishments and rewards required.

Opinions may vary as to the complexity which is suitable in the child machine. One might try to make it as simple as possible consistent with the general principles. Alternatively one might have a complete system of logical inference "built in."² In the latter case the store would be largely occupied with definitions and propositions. The propositions would have various kinds of status, *e.g.*, well-established facts, conjectures, mathematically proved theorems, statements given by an authority, expressions having the logical form of proposition but not belief-value. Certain propositions may be described as "imperatives." The machine should be so constructed that as soon as an imperative is classed as "well established" the appropriate action automatically takes place. To illustrate this, suppose the teacher says to the machine, "Do your homework now." This may cause "Teacher says 'Do your homework now'" to be included amongst the well-established facts. Another such fact might be "Everything that teacher says is true." Combining these may eventually lead to the imperative, "Do your homework now," being included amongst the well-established facts, and this, by the construction of the machine, will mean that the homework actually gets started, but the effect is very satisfactory. The processes of inference used by the machine need not be such as would satisfy the most exacting logicians. There might for instance be no hierarchy of types. But this need not mean that type fallacies

will occur, any more than we are bound to fall over unfenced cliffs. Suitable imperatives (expressed *within* the systems, not forming part of the rules of the system) such as "Do not use a class unless it is a subclass of one which has been mentioned by teacher" can have a similar effect to "Do not go too near the edge."

The imperatives that can be obeyed by a machine that has no limbs are bound to be of a rather intellectual character, as in the example (doing homework) given above. Important amongst such imperatives will be ones which regulate the order in which the rules of the logical system concerned are to be applied. For at each stage when one is using a logical system, there is a very large number of alternative steps, any of which one is permitted to apply, so far as obedience to the rules of the logical system is concerned. These choices make the difference between a brilliant and a footling reasoner, not the difference between a sound and a fallacious one. Propositions leading to imperatives of this kind might be "When Socrates is mentioned, use the syllogism in Barbara" or "If one method has been proved to be quicker than another, do not use the slower method." Some of these may be "given by authority," but others may be produced by the machine itself, *e.g.* by scientific induction.

The idea of a learning machine may appear paradoxical to some readers. How can the rules of operation of the machine change? They should describe completely how the machine will react whatever its history might be, whatever changes it might undergo. The rules are thus quite time-invariant. This is quite true. The explanation of the paradox is that the rules which get changed in the learn-

² Or rather "programmed in" for our child machine will be programmed in a digital computer. But the logical system will not have to be learnt.

ing process are of a rather less pretentious kind, claiming only an ephemeral validity. The reader may draw a parallel with the Constitution of the United States.

An important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil's behavior. This should apply most strongly to the later education of a machine arising from a child machine of well-tried design (or programme). This is in clear contrast with normal procedure when using a machine to do computations: one's object is then to have a clear mental picture of the state of the machine at each moment in the computation. This object can only be achieved with a struggle. The view that "the machine can only do what we know how to order it to do,"³ appears strange in face of this. Most of the programmes which we can put into the machine will result in its doing something that we cannot make sense of at all, or which we regard as completely random behaviour. Intelligent behaviour presumably consists in a departure from the completely disciplined behaviour involved in computation, but a rather slight one, which does not give rise to random behaviour, or to pointless repetitive loops. Another important result of preparing our machine for its part in the imitation game by a process of teaching and learning is that "human fallibility" is likely to be omitted in a rather natural way, *i.e.*, without special "coaching." Processes that are learnt do not produce a hundred per cent cer-

tainty of result; if they did they could not be unlearned.

It is probably wise to include a random element in a learning machine. A random element is rather useful when we are searching for a solution of some problem. Suppose for instance we wanted to find a number between 50 and 200 which was equal to the square of the sum of its digits, we might start at 51 then try 52 and go on until we got a number that worked. Alternatively we might choose numbers at random until we got a good one. This method has the advantage that it is unnecessary to keep track of the values that have been tried, but the disadvantage that one may try the same one twice, but this is not very important if there are several solutions. The systematic method has the disadvantage that there may be an enormous block without any solutions in the region which has to be investigated first. Now the learning process may be regarded as a search for a form of behaviour which will satisfy the teacher (or some other criterion). Since there is probably a very large number of satisfactory solutions the random method seems to be better than the systematic. It should be noticed that it is used in the analogous process of evolution. But there the systematic method is not possible. How could one keep track of the different genetical combinations that had been tried, so as to avoid trying them again?

We may hope that machines will eventually compete with men in all purely intellectual fields. But which are the best ones to start with? Even this is a difficult decision. Many people think that a very abstract activity, like the playing of chess, would be best. It can also be maintained that it is best to provide the machine with the

³ Compare Lady Lovelace's statement which does not contain the word "only."

best sense organs that money can buy, and then teach it to understand and speak English. This process could follow the normal teaching of a child. Things would be pointed out and named, etc. Again I do not know what the right answer is, but I think both approaches should be tried.

We can only see a short distance ahead, but we can see plenty there that needs to be done.

Bibliography

- Samuel Butler, *Erewhon*, London, 1865.
 Chapters 23, 24, 25, *The Book of the Machines*.
- Alonzo Church, "An Unsolvability Problem of Elementary Number Theory," *American J. of Math.*, 58 (1936), 345-363.
- K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I," *Monatshefte für Math. und Phys.* (1931), 173-189.
- D. R. Hartree, *Calculating Instruments and Machines*, New York, 1949.
- G. Jefferson, "The Mind of Mechanical Man." Lister Oration for 1949. *British Medical Journal*, vol. i (1949), 1105-1121.
- S. C. Kleene, "General Recursive Functions of Natural Numbers," *American J. of Math.*, 57 (1935), 153-173 and 219-244.
- Countess of Lovelace, "Translator's notes to an article on Babbage's Analytical Engine," *Scientific Memoirs* (ed. by R. Taylor), vol. 3 (1842), 691-731.
- Bertrand Russell, *History of Western Philosophy*, London, 1940.
- A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proc. London Math. Soc.* (2), 42 (1937), 230-265.

INTRODUCTION

In this article, the first chapter of sixteen in Tinbergen's fascinating book, *Curious Naturalists*, you will read a very human description of the ethological approach to understanding behavior. There are now a fair number of ethologists whose work includes elaborate field studies followed by specific experimentation—either in the field or in the laboratory—to test conjectures about the mechanisms of behavior which are at work. Here Tinbergen describes his studies of the bee-hunting wasp, *Philanthus*. This tiny creature with a literally microscopic brain and very primitive compound eyes displays a remarkable ability to find his way about visually in the natural world and to learn very rapidly how to make use of new landmarks when they become available.

You might think about the following questions while reading this article. Do you think the wasp plans his food gathering expeditions in the same way as a man might plan a hunting trip? If not, what kinds of mechanisms do you suppose are at work? Do you think any of these ever play a role in a human's existence? How do you suppose the bee hunter is able to recognize a bee? Why don't they ever bring home other kinds of prey? Suppose that you were watching a given burrow and saw what seemed to be a single wasp going in, flying away, and then returning after a few minutes. How could you find out if it was really the same wasp or whether perhaps several different wasps were coming and going? Do you think that the wasps depend on any reinforcement (like the food pellets given to rats in laboratory experiments) in order to learn the position of new landmarks around their nests?

Tinbergen now lives in England and occasionally uses words which may not be familiar to the American reader. The Chanterelle referred to is an edible mushroom! The crofter is a tenant farmer.

If you enjoy this chapter as much as we have, you may wish to buy the paperback, *Curious Naturalists*, published in the Natural History Library by Doubleday and Company.

THE BEE-HUNTERS OF HULSHORST

Nicolaas Tinbergen

As a child I thoroughly disliked insects and always suspected that, apart from crawling over one's skin, they all might bite or sting. Later, however, I spent many summers studying the habits of one of the fiercest of digger wasps, *Philanthus*, the bee-killer, and gradually I grew very fond of this insect and of many others. I owe this interest to the lovely country of Hulshorst.

In my early twenties, when I had been a zoology student for several years, I still showed a certain pre-occupation with hockey, pole-jumping, skating, camping and bird-photography. My zoology professor did not altogether approve of this; in fact, he had given me up as an incipient zoologist, and I cannot altogether blame him. Yet I was a budding naturalist—of a kind—and without being aware of it I was waiting for the critical stimulus that would start me on what was to become my hobby and my job.

The summer vacation usually found my parents, with all their children and many friends, in occupation of a small

cottage in one of the less densely populated areas of crowded little Holland, amid wide expanses of glacial sands which, fortunately enough, were good for nothing except, on the best parts, growing low-quality timber. Some irregular strips of fine deciduous trees such as Beech and Oak bordered a brook that wound its way through the heaths and sand dunes on its way to the Zuiderzee. Several square miles of the hilly sands had been planted with Scots Pine, but large stretches were covered with a carpet of Heather, and other parts again were just bare sand of the poorest sort. Isolated wind-blown Pines, and little dunes with a sparse vegetation of Marram Grass, broke the monotony of these arid plains.

But don't think that this country was dull—far from it! On countless long walks, at first in the summer, later in all seasons of the year, we discovered its many hidden treasures. The cool green Beech wood bordering the brook harboured such fine birds as Black Woodpeckers, Honey Buzzards and Woodcock. Quietly following the little stream, we often surprised Roe-deer nibbling at fresh green shoots, or drinking in a quiet corner. In the autumn, when the fallen Beech leaves glowed like polished copper, we were delighted to find the most wonderful

Reprinted with permission from Curious Naturalists by Nicolaas Tinbergen. Basic Books, Doubleday Series, Museum of Natural History Books, 1958, pp. 1-18, and the Hamlyn Publishing Group Limited. Figure 109 reprinted from The Study of Instinct by N. Tinbergen by permission of Clarendon Press, Oxford, England.

mushrooms; collecting the edible species gave a culinary value to our leisurely walks, and made us sympathize with the Red Squirrels who were busy nibbling at acorns or burying them.

The Scots Pine plantations, at first glance so monotonous, were full of fascinating creatures, many of them beautifully camouflaged. The larvae of the Pine Hawk Moths and several other caterpillars were living among the needles, eating away steadily, blending with their background through their wonderful pattern of green and white stripes. In the poorer parts of the woods, where the trunks were covered with lichens, a peculiar fauna of spiders, bugs and moths was found, all matching their background by their remarkably efficient colour patterns of a pale lichen green covered with black blotches.

On the greenish grey carpet of reindeer moss, soft as velvet or crisp as toast, according to the weather, the lovely orange-ochre Chanterelles glowed all through the summer and autumn—and they supplied us with many a good meal.

Flocks of tits travelled through the tops of the trees, continually whispering to keep in touch with each other. Their whispers stopped only when a sharp 'zee' of one of them warned them of an attacking Sparrow Hawk. These, as well as their bigger cousins, the Goshawks, could be found nesting in the older parts of the woods.

One dramatic winter day found us walking through these woods during a fierce blizzard. Starting as rain, then turning to sleet, then, with dropping temperature, into dry snow, huge masses of moisture were hurled down on the wood, stuck to the needles and the twigs, and finally, through their accumulated weight, began to break even the strongest branches. Above the

roar of the gale, the cracking and snapping of the overloaded branches rang through the woods like so many rifle shots. Looking at this, not with the eyes of a forester but with those of the naturalist, we thought ourselves fortunate in witnessing one of those rare, critical days, so interesting to the ecologist because they may well have tremendous consequences for wild life.

Quite different were the wide heaths. In late summer, when the Heather bore its myriads of flowers, there was a purple glow over the undulating plains, fading away in the heat haze, the horizon shimmering and quivering in the hot sun. The flowers of the Heather, producing a delicious nectar, attracted innumerable insects, whose continuous buzzing was as much part of the scene as the sound of the surf is part of the sea shore. Here and there we would find rows of primitive straw bee-hives, belonging to small crofters living up to twenty miles away, who had carted them up for the short but prolific season.

On long walks one could suddenly chance upon tiny moors, with cotton grass and many colourful plants; a family of Curlews disturbed by our intrusion would fly off, uttering their melodious calls.

Finally, even the bare sandy stretches had a charm of their own. The sparse vegetation could not manage to cover the soil entirely, and on windy days the sand was blown over the dune crests. Thus, the small hills were continually on the move, and since south-westerly winds prevail in this part of the world, the bulk of the sand was always shifting to the north-east. On the south-westerly edge of the sands there were flat, blown-out plains, whereas on the north-east side the bare sand tongues moved slowly but irresistibly

forward, burying everything in their path, heath and forest alike.

On the flats the gravel, too heavy to be blown away, was left behind, and this put a stop to wind erosion. Lichens, mosses, grasses and finally Pine seedlings settled here, and gradually such plains were covered by a very open vegetation of Pine, *Calluna*, and scattered patches of lichens and mosses.

The sands, and these natural Pine woods, were not as poor as they might seem. They were lovely insect country. Thousands of Ant Lions preyed on the Red Wood Ants which abounded where isolated Oak and Birch trees supported a rich supply of insects. Pine Hawk Moths lived on the isolated Pines. Nightjars were abundant. Dung Beetles lived on the droppings of the many Rabbits; on sunny days they flew about high up in the sky, and Hobbies, those elegant little falcons, spent hours hawking them on the wing. Tiger Beetles, large predatory flies (*Asilus*) and many kinds of burrowing wasps earned a living by preying on other insects, either those living there on the sparse vegetation or others, such as the hapless Crane Flies, that strayed into the area.

We lived in that country long enough to have seen it in a great variety of circumstances and moods, and gradually every part of it acquired significance; the whole area became charged with our experiences, many of which I shall never forget. Here we had found the flint implements left by people living in this place several thousand years ago. There we had seen the lightning strike a tree, shattering its bark, and throwing large pieces of it all around. At another place, right in the middle of the sands, we had found another witness of lightning, where it had struck the ground—a 'ful-

gurite' or tube of molten sand. We dug the whole thing up, all 12 ft. of it, and it is now in the Leiden Geological Museum.

Some parts of the woods are forever imprinted in our memory as the haunts of the Goshawk; we spent many hours in a tree-top hide watching this formidable predator feeding its young in the large nest. And one particular corner of the sands, a blown-out, flat valley of a quarter of a mile square, is my favourite spot since it was there that I discovered *Philanthus*, the digger wasp that made me take up my first experimental study—and so helped me to gain some self-respect.

On a sunny day in the summer of 1929 I was walking rather aimlessly over the sands, brooding and a little worried. I had just done my finals, had got a half-time job, and was hoping to start on research for a doctor's thesis. I wanted very much to work on some problem of animal behaviour and had for that reason rejected some suggestions of my well-meaning supervisor. But rejecting sound advice and taking one's own decisions are two very different things, and so far I had been unable to make up my mind.

While walking about, my eye was caught by a bright orange-yellow wasp the size of the ordinary jam-loving *Vespa*. It was busying itself in a strange way on the bare sand. With brisk, jerky movements it was walking slowly backwards, kicking the sand behind it as it proceeded. The sand flew away with every jerk. I was sure that this was a digger wasp. The only kind of that size I knew was *Bembex*, the large fly-killer. But this was no *Bembex*. I stopped to watch it, and soon saw that it was shovelling sand out of a burrow. After ten minutes of this, it turned round, and now, facing

away from the entrance, began to rake loose sand over it. In a minute the entrance was completely covered. Then the wasp flew up, circled a few times round the spot, describing wider and wider loops in the air, and finally flew off. Knowing something of the way of digger wasps, I expected it to return with a prey within a reasonable time, and decided to wait.

Sitting down on the sand, I looked round and saw that I had blundered into what seemed to be a veritable wasp town. Within ten yards I saw more than twenty wasps occupied at their burrows. Each burrow had a patch of yellow sand round it the size of a hand, and to judge from the number of these sand patches there must have been hundreds of burrows.

I had not to wait long before I saw a wasp coming home. It descended slowly from the sky, alighting after the manner of a helicopter on a sand patch. Then I saw that it was carrying a load, a dark object about its own size. Without losing hold of it, the wasp made a few raking movements with its front legs, the entrance became visible and, dragging its load after it, the wasp slipped into the hole.

At the next opportunity I robbed a wasp of its prey, by scaring it on its arrival, so that it dropped its burden. Then I saw that the prey was a Honey Bee.

I watched these wasps at work all through that afternoon, and soon became absorbed in finding out exactly what was happening in this busy insect town. It seemed that the wasps were spending part of their time working at their burrows. Judging from the amount of sand excavated these must have been quite deep. Now and then a wasp would fly out and, after half an hour or longer, return with a load, which was then dragged in. Every time

I examined the prey, it was a Honey Bee. No doubt they captured all these bees on the heath for all the to and fro traffic was in the direction of the south-east, where I knew the nearest heath to be. A rough calculation showed that something was going on here that would not please the owners of the bee-hives on the heath; on a sunny day like this several thousand bees fell victims to this large colony of killers!

As I was watching the wasps, I began to realize that here was a wonderful opportunity for doing exactly the kind of field work I would like to do. Here were many hundreds of digger wasps—exactly which species I did not know yet, but that would not be difficult to find out. I had little doubt that each wasp was returning regularly to its own burrow, which showed that they must have excellent powers of homing. How did they manage to find their way back to their own burrow?

At that time the excellent studies of several German zoologists, notably E. Wolf, had already proved that Honey Bees were quite good at homing, and much had already been discovered of the way they did it. But I knew that the homing abilities of solitary wasps were still unexplored, and that the observations of that master of entomology, Henry Fabre and of his followers, such as Ferton, Rau and others, rather suggested that such wasps had quite mysterious powers of finding their way back to their own burrows. But the work of these naturalists, though admirable in many ways, had always left me unsatisfied, and obviously something had to be done about this problem. Here seemed to be a first-rate opportunity. I then and there decided to tackle it.

There was another puzzling thing about these wasps. If they specialized

so entirely on Honey Bees—as they seemed to do—how did they recognize them among the teeming thousands of insects of so many kinds that were feasting on the Heather? Would it be possible to find out?

My worries were over; I knew what I wanted to do. This day, as it turned out, was a milestone in my life. For several years to come I was to spend my summers with these wasps, first alone, then with an ever-growing group of co-workers, most of them zoology students of the University of Leiden. Soon we began to study other insects as well, and thus started what has come to be a tradition in our Zoology Laboratory, of organizing biological summer camps year after year at Hulshorst. We owe a great debt of gratitude to the generous owner of this lovely stretch of country, Mr. A. E. Jurriaanse, who so cordially welcomed us every summer, who allowed us to go about our business (which must at times have seemed a little odd), and who gave us support of so many kinds.

When in 1949 I left Holland to settle in Oxford my colleague, Dr. Jan van Iersel, took over and, while I write this, the Hulshorst work is still going strong on a considerably expanded scale.

Settling down to work, I started spending the wasps' working days (which lasted from about 8 a.m. till 6 p.m. and so did not put too much of a strain on me) on the 'Philanthus plains', as we called this part of the sands as soon as we had found out that *Philanthus triangulum* Fabr. was the official name of this bee-killing digger wasp. Its vernacular name was 'Bee-Wolf'.

An old chair, field glasses, notebooks, and food and water for the day

were my equipment. The local climate of the open sands was quite amazing, considering that ours is a temperate climate. Surface temperatures of 110° F were not rare, and judging from the response of my skin, which developed a dark tan, I got my share of ultra-violet radiation.

My first job was to find out whether each wasp was really limited to one burrow, as I suspected from the unhesitating way in which the home-coming wasps alighted on the sand patches in front of the burrows. I installed myself in a densely populated quarter of the colony, five yards or so from a group of about twenty-five nests. Each burrow was marked and mapped. Whenever I saw a wasp at work at a burrow, I caught it and, after a short unequal struggle, adorned its back with one or two colour dots (using quickly drying enamel paint) and released it. Such wasps soon returned to work, and after a few hours I had ten wasps, each marked with a different combination of colours, working right in front of me. It was remarkable how this simple trick of marking my wasps changed my whole attitude to them. From members of the species *Philanthus triangulum* they were transformed into personal acquaintances, whose lives from that very moment became affairs of the most personal interest and concern to me.

While waiting for events to develop, I spent my time having a close look at the wasps. A pair of lenses mounted on a frame that could be worn as spectacles enabled me, by crawling up slowly to a working wasp, to observe it, much enlarged, from a few inches away. When seen under such circumstances most insects reveal a marvellous beauty, totally unexpected as long as you observe them with the unaided eye. Through my lenses I could look

at my *Philanthus* right into their huge compound eyes; I saw their enormous, claw-like jaws which they used for crumbling up the sandy crust; I saw their agile black antennae in continuous, restless movement; I watched their yellow, bristled legs rake away the loose sand with such vigour that it flew through the air in rhythmic puffs, landing several inches behind them.

Soon several of my marked wasps stopped working at their burrows, raked loose sand back over the entrance, and flew off. The take-off was often spectacular. Before leaving they circled a little while over the burrow, at first low above the ground, soon higher, describing ever widening loops; then flew away, but returned to cruise once more low over the nest. Finally, they would set out in a bee-line, fifteen to thirty feet above the ground, a rapidly vanishing speck against the blue sky. All the wasps disappeared towards the south-east. Half a mile away in that direction the bare sands bordered upon an extensive heath area, buzzing with bees. This, as I was to see later, was the wasps' hunting area.

The curious loops my wasps described in the air before leaving their home area had been described by other observers of many other digger wasps. Philip Rau had given them the name of 'locality studies'. Yet so far nobody had proved that they deserved that name; that the wasps actually took in the features of the burrow's surroundings while circling above them. To check this if possible was one of my aims—I thought that it was most probable that the wasps would use landmarks, and that this locality study was what the name implied. First, however, I had to make sure that

my marked wasps would return to their own holes.

When my wasps had left, there began one of those periods of patient waiting which are usual in this kind of work. It was, of course, necessary to be continually on the look-out for returning wasps; at the same time it was tempting to look round and watch the multitude of other creatures that were busy on the hot plains. For I soon discovered that I had many neighbours. First of all, there were other diggers about. Among the *Philanthus* burrows there were some that looked different—the sand patches were a little larger and less regular. These belonged to the largest of our digger wasps, the fly-killing *Bembex*, that is almost the size of a Hornet. Buzzing loudly, flying with terrific speed low over the ground, these formidable wasps dashed to and fro, and it took me a long time before I saw more of them than a momentary sulphur-yellow flash. A Leafcutter Bee was coming home carrying its 'wall paper', a neat circular disc cut out of a rose leaf. Its burrow, scarcely visible, was in the carpet of dry moss just beyond the *Philanthus* settlement. Robber Flies (*Asilus crabroniformis*) whizzed past, catching flies and other insects in the air. Sometimes they would make a mistake and attack *Philanthus*. A brief but furious buzzing struggle, and the two fell apart, *Asilus* darting off to find a less petulant prey, *Philanthus* returning to its burrow. Small grasshoppers came walking by, greedy and single-minded, devouring one grass shoot after the other, or spending hours chirping their song, or courting, with pathetic perseverance, seemingly unconcerned females.

On some days we saw endless processions of migrating Cabbage White but-

terflies crossing the plains in a continuous stream of scattered formations, usually going north-west. Hobbies came and levied their toll on them, as well as on the many dragonflies and the clumsy Dung Beetles. Bumblebees often zoomed past on their long, mysterious trips and, like the Cabbage Whites, would show their interest in nectar-filled blue flowers by alighting on anything blue in our equipment.

In August the monotony of the deep blue sky might be broken by a lonely Osprey, flying in from his fishing grounds north of us—the coastal waters of the Zuiderzee—to settle in the crown of an old Pine, there to dream away the hours of digestion. Or a group of lovely, black-and-white Storks, on migration to Africa, might come sailing by, stopping in the rising air over the hot sands to soar up in wide circles, higher and higher, until they resumed their glide south in search of the next 'thermal'. Thus sitting and waiting for the wasps was never dull, if only one kept one's eyes open.

To return to my marked wasps: before the first day was over, each of them had returned with a bee; some had returned twice or even three times. At the end of that day it was clear that each of them had its own nest, to which it returned regularly.

On subsequent days I extended these observations and found out some more facts about the wasps' daily life. As in other species, the digging of the large burrows and the capturing of prey that served as food for the larvae was exclusively the task of the females. And a formidable task it was. The wasps spent hours digging the long shafts, and throwing the sand out. Often they stayed down for a time and, waiting for them to re-

appear, my patience was often put to a hard test. Eventually, however, there would be some almost imperceptible movement in the sand, and a small mound of damp soil was gradually lifted up, little by little, as if a miniature Mole were at work. Soon the wasp emerged, tail first, and all covered with sand. One quick shake, accompanied by a sharp staccato buzz, and the wasp was clean. Then it began to mop up, working as if possessed, shovelling the sand several inches away from the entrance.

I often tried to dig up the burrows to see their inner structure. Usually the sand crumbled and I lost track of the passage before I was ten inches down, but sometimes, by gently probing with a grass shoot first, and then digging down along it, I succeeded in getting down to the cells. These were found opening into the far end of the shaft, which itself was a narrow tube, often more than 2 ft. long. Each cell contained an egg or a larva with a couple of Honey Bees, its food store. A burrow contained from one to five cells. Each larva had its own living room-cum-larder in the house, provided by the hard-working female. From the varying number of cells I found in the nests, and the varying ages of the larvae in one burrow, I concluded that the female usually filled each cell with bees before she started to dig a new cell, and I assumed that it was the tunnelling out of a new cell that made her stay down for such long spells.

I did not spend much time digging up the burrows, for I wanted to observe the wasps while they were undisturbed. Now that I was certain that each wasp returned regularly to her own burrow, I was faced with the problem of her orientation. The en-

tire valley was littered with the yellow sand patches; how could a wasp, after a hunting trip of about a mile in all, find exactly her own burrow?

Having seen the wasps make their 'locality studies', I naturally believed that each female actually did what this term implied: take her bearings. A simple test suggested that this was correct. While a wasp was away I brushed over the ground surrounding the nest entrance, moving all possible landmarks such as pebbles, twigs, tufts of grass, Pine cones, etc., so that over an area of 3-4 square metres none of them remained in exactly the same place as before. The burrow itself, however, I left intact. Then I awaited the wasp's return. When she came, slowly descending from the skies, carrying her bee, her behavior was striking. All went well until she was about 4 ft. above the ground. There she suddenly stopped, dashed back and forth as if in panic, hung motionless in the air for a while, then flew back and up in a wide loop, came slowly down again in the same way, and again shied at the same distance from the nest. Obviously she was severely disturbed. Since I had left the nest itself, its entrance, and the sand patch in front of it untouched, this showed that the wasp was affected by the change in the surroundings.

Gradually she calmed down, and began to search low over the disturbed area. But she seemed to be unable to find the nest. She alighted now here, now there, and began to dig tentatively at a variety of places at the approximate site of the nest entrance. After a while she dropped her bee and started a thorough trial-and-error search. After twenty-five minutes or so she stumbled on the nest entrance as if by accident, and only then did she take up her bee and drag it in.

A few minutes later she came out again, closed the entrance, and set off. And now she had a nice surprise in store for me: upon leaving she made an excessively long 'locality study': for fully two minutes she circled and circled, coming back again and again to fly over the disturbed area before she finally zoomed off.

I waited for another hour and a half, and had the satisfaction of seeing her return once more. And what I had hoped for actually happened: there was scarcely a trace of hesitation this time. Not only had the wasp lost her shyness of the disturbed soil, but she now knew her way home perfectly well.

I repeated this test with a number of wasps, and their reactions to my interference were roughly the same each time. It seemed probable, therefore, that the wasps found their way home by using something like landmarks in the environment, and not by responding to some stimulus (visual or otherwise) sent out by the nest itself. I had now to test more critically whether this was actually the case.

The test I did next was again quite simple. If a wasp used landmarks it should be possible to do more than merely disturb her by throwing her beacons all over the place; I ought to be able to mislead her, to make her go to the wrong place, by moving the whole constellation of her landmarks over a certain distance. I did this at a few nests that were situated on bare sandy soil and that had only a few, but conspicuous, objects nearby, such as twigs, or tufts of grass. After the owner of such a nest was gone, I moved these two or three objects a foot to the south-west, roughly at right angles to the expected line of approach. The result was as I had hoped for and expected, and yet I could not

help being surprised as well as delighted: each wasp missed her own nest, and alighted at exactly the spot where the nest 'ought' to be according to the landmarks' new positions! I could vary my tests by very cautiously shooing the wasp away, then moving the beacons a foot in another direction, and allowing the wasp to alight again. In whatever position I put the beacons, the wasp would follow them. At the end of such a series of tests I replaced the landmarks in their original position, and this finally enabled the wasp to return to her home. Thus the tests always had a happy ending—for both of us. This was no pure altruism on my part—I could now use the wasp for another test if I wished.

When engaged in such work, it is always worth observing oneself as well as the animals, and to do it as critically and as detachedly as possible—which, of course, is a tall order. I have often wondered why the outcome of such a test delighted me so much. A rationalist would probably like to assume that it was the increased predictability resulting from the test. This was a factor of considerable importance, I am sure. But a more important factor still (not only to me, but to many other people I have watched in this situation) is of a less dignified type: people enjoy, they relish the satisfaction of their desire for power. The truth of this was obvious, for instance, in people who enjoyed seeing the wasps being misled without caring much for the intellectual question whether they used landmarks or not. I am further convinced that even the joy of gaining insight was not often very pure either; it was mixed with pride at having had success with the tests.

To return to the wasps: next I tried to make the wasps use landmarks

which I provided. This was not only for the purpose of satisfying my lust for power, but also for nobler purposes, as I hope to show later. Since changing the environment while the wasp was away disturbed her upon her return and even might prevent her from finding her nest altogether, I waited until a wasp had gone down into her nest, and then put my own landmarks round the entrance—sixteen Pine cones arranged in a circle of about eight inches diameter.

The first wasp to emerge was a little upset, and made a rather long locality study. On her return home, she hesitated for some time, but eventually alighted at the nest. When next she went out she made a really thorough locality study, and from then on everything went smoothly. Other wasps behaved in much the same way, and next day regular work was going on at five burrows so treated. I now subjected all five wasps, one by one, to a displacement test similar to those already described. The results, however, were not clear-cut. Some wasps, upon returning, followed the cones; but others were not fooled, and went straight home, completely ignoring my beacons. Others again seemed to be unable to make up their minds, and oscillated between the real nest and the ring of cones. This half-hearted behaviour did not disturb me, however, for if my idea was correct—that the wasps use landmarks—one would rather expect that my tests put the wasps in a kind of conflict situation: the natural landmarks which they must have been using before I gave them the Pine cones were still in their original position; only the cones had been moved. And while the cones were very conspicuous landmarks, they had been there for no more than one day. I therefore put all the cone-rings back

and waited for two more days before testing the wasps again. And sure enough, this time the tests gave a hundred per cent preference for the Pine cones; I had made the wasps train themselves to my landmarks.

The rest of this first summer I spent mainly in consolidating this result in various ways. There was not much time to do this, for the season lasts only two months; by the end of August the wasps became sluggish, and soon after they died, leaving the destiny of their race in the hands of the pupae deep down in the sand, which were to lie there dormant until next July. And even in this short summer season the wasps could not work steadily, but were active on dry sunny days only—and of these a Dutch summer rarely supplies more than about twenty in all.

However, I had time to make sure that the wasps relied for their homing

mainly on vision. First, I could cut off their antennae—the bearers of delicate organs of smell, of touch and of other sense organs—without at all disturbing their orientation. Second, when, in other tests, I covered the eyes of intact wasps with black paint, the wasps could not fly at all. Removing the cover of paint restored their eyesight, and with it their normal behaviour. Furthermore, when I trained a wasp to accept a circle of Pine cones together with two small squares of cardboard drenched in Pine oil, which gave off a strong scent, displacement of the cones would mislead the wasps in the usual way, but moving the scented squares had not the slightest effect. Finally, when wasps used to rings of cones were given, instead of cones, a ring of grey pebbles a foot from the nest, they followed these pebbles. This can only have been due to the pebbles being visually similar to the cones.

This was about as far as I got in that first summer. I had had my first thrills as an experimenter. However simple my observations and tests had been, they had led me to some real discoveries. I had tasted that particular brand of triumph that is the reward of true exploration. From now on my plans for the summer vacations for years to come were made: I had to keep going back to Hulshorst and find out more about the bee-killers. As it turned out, this decision fixed the summer plans of my wife, of all our children, and of several 'generations' of students as well: for the field work on the sandy plains went on for many years, even after the bee-killers, unusually abundant in these first years, had fallen back to their normal status of a relatively rare species, and so forced us to abandon their study.

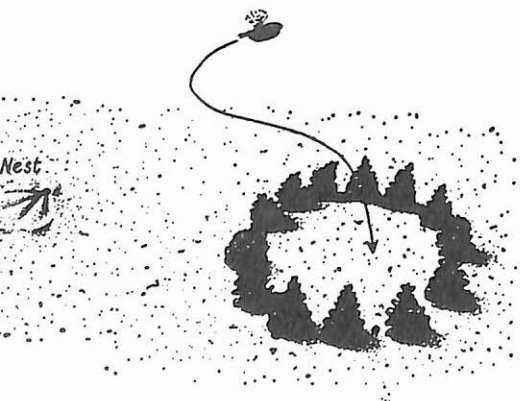


Figure 109

A homing test. *Philanthus* returns to the displaced pine cones and fails to find her burrow.

INTRODUCTION

There has long been a movement in psychology to explain all overt behavior as simple stimulus-response (S-R) associations. As a result of the past history of an organism, given stimuli are automatically followed by certain responses and behavior patterns are completely determined by these simple connections.

In the brief section which follows, Watson outlines the extreme point of view that thinking is essentially subvocal speech—simply a low intensity form of talking. Accordingly, thinking, like talking, and therefore like any other overt response should be thought of as simple stimulus-response association. As you read this remember that it is an extreme example of the connectionist point of view—so extreme as to appear almost as a caricature today. But the general idea that all behavior, overt and covert, should be thought of as stimulus-response associations is still a very prominent idea in modern psychology. As you read Watson's section, try to reflect not only on this specific argument but also on the more general stimulus-response conception of behavior.

WHAT IS THINKING?

John B. Watson

Before trying to understand the behaviorists' theory of thinking, won't the reader please pick up any introspective psychological text and read the chapter on thinking? Won't he then try to digest some of the pabulum the philosophers have offered us on this all-important function? I have tried to understand it. I had to give it up. I believe the reader will give it up too. But until he has read their explanations, don't quarrel with the behaviorist for weaknesses in his presentation. His own theory is quite simple. The only difficulty about it lies in our previous organization. We begin to resist it, to show negative reactions, the moment we hear about it. We have been trained both at our mother's knee and in psychological laboratories to say that thinking is something peculiarly uncorporeal, something very intangible, very evanescent, something peculiarly mental. To the behaviorist this resistance is due to the reluctance of the psychologists to give up the teaching of religion in their psychology. Thinking, on account of the concealed nature of

the musculature with which it is done, has always been inaccessible to unaided observation and to direct experimentation. And there is always a strong inclination to attach a mystery to something one can't see. As new scientific facts are discovered we have fewer and fewer phenomena which cannot be observed, hence fewer and fewer pegs upon which to hang folklore. The behaviorist advances a natural science theory about thinking which makes it just as simple, and just as much a part of biological processes, as tennis playing.

The Behaviorist's View of Thought

The behaviorist advances the view that *what the psychologists have hitherto called thought is in short nothing but talking to ourselves*. The evidence for this view is admittedly largely theoretical but it is the one theory so far advanced which explains thought in terms of natural science. I wish here expressly to affirm that in developing this view I have never believed that the *laryngeal movement* as such played the predominating rôle in thought. I admit that in my former presentations I have, in order to gain pedagogical simplicity, expressed myself in ways which can be so interpreted. We have all had the proofs before us time and again that

Reprinted with permission from Behaviorism by John B. Watson, W. W. Norton & Company, Inc. Copyright 1924, 1925 by the Peoples Institute Publishing Co., Inc.; Copyright renewed 1952 by John B. Watson. Revised edition Copyright 1930 by W. W. Norton & Company, Inc. Copyright renewed 1958 by John B. Watson, pp. 237-40.

the larynx can be removed without completely destroying a person's ability to think. Removal of the larynx does destroy articulate speech but it does not destroy whispered speech. Whispered speech (without articulation) depends upon muscular responses of the cheek, tongue, throat and chest—organization which, to be sure, has been built up with the use of the larynx, but which remains ready to function after the larynx has been removed. Any one who has read my various presentations knows that I have tried everywhere to emphasize the enormous complexity of the musculature in the throat and chest. To claim that a mass of cartilage such as that composing the larynx is responsible for thought (internal speech) is like saying that the bone and cartilage that make up the elbow joint form the chief organ with which one plays tennis.

My theory does hold that the muscular habits learned in overt speech are responsible for implicit or internal speech (thought). It holds, too, that there are hundreds of muscular combinations with which one can say either aloud or to himself almost any word, so rich and so flexible is language organization and so varied are our overt speech habits. A good imitator, we know, can say the same phrases in dozens of different ways, in a bass voice, in a tenor voice, in a mezzo, in a soprano, in a loud whisper, in a soft whisper, as an English cockney would say them, as a broken-English-speaking Frenchman would say them, as a Southerner would say

them, as a child. The number and variety of habits we form in the speaking of almost every word is thus well-nigh legion. We use speech, from infancy on, a thousand times to using our hands once. From this circumstance there grows up a complexity of organization which even the psychologist seemingly cannot grasp. Again, after our overt speech habits are formed, we are constantly talking to ourselves (thought). New combinations occur, new complexities arise, new substitutions take place—for example, where the shrug of the shoulders or the movement of any other bodily part becomes substituted for a word. Soon, *any, and every* bodily response *may become a word substitute*.

The alternative sometimes advanced to this theory is that so-called central processes may take place in the brain so faintly that no neural impulse passes out over the motor nerve to the muscle, hence no response takes place in the muscles and glands. Even Lashley and his students, on account of their strong interest in the nervous system, seem to hold this view. Recently Agnes M. Thorson has found that tongue movements are not universally present in internal speech. This, even if true, can have no bearing upon the present view. The tongue, while bearing very delicate receptors, is on the muscular side a bulk organ for rolling our food around. It plays a part in internal speech to be sure, but probably about the same part that the fist of the jazz cornet player plays when he thrusts it into the horn to modify the sound.

INTRODUCTION

A small number of human activities seem to be readily characterized as simple, quickly occurring responses elicited by an equally simple stimulus. The human eye blink is one such case; but, by far the greater proportion of human activities seems more readily thought of in terms of an organized series of responses. For example, there is an elaborate and integrated series of tongue, lip, and jaw movements involved in the simple pronunciation of the word "the." Other sequences are involved in overt behaviors such as taking a step, hitting a nail with a hammer, or balancing an orange on your nose. The traditional psychological model for such series of movements or actions is that of a response chain: one response occurs and then its occurrence acts as a stimulus for the next response to occur, which in turn acts as a stimulus for the following one, etc. In certain situations this model may be appropriate. A painter may make a stroke on his canvas, sit back to look at the whole, make another stroke, sit back to evaluate, make another stroke, etc. Here each stroke is, in some sense, a stimulus for the stroke which follows it. But the essential point made in the following article is that this simple mechanism cannot possibly account for the complexity of human behavior. You should enjoy Lashley's clever analyses in support of his assertion. This article was addressed to a group of physiologists, since physiologists had long accepted the oversimplified psychological model mentioned; but psychologists, too, have learned (or *should* have learned) a great deal from this work.

As you read think about your own activities in any ten minute period. How many of them are organized into a sequence of movements, units, or steps which are executed one after another either very rapidly or more slowly?

In order to appreciate the difference between chained S-R connections, and the sort of mechanism postulated by Lashley, imagine the problems in rapidly delivering a long sequence of instructions to a very distant space ship. The instructions would be sent out sequentially, one after another. Because of the great time needed for the instructions to reach the rocket and for communication to travel from the rocket to earth, it would be completely hopeless to give one instruction, wait to see that

it had been properly received, and then give another. Instructions must be issued consecutively in a long sequence without waiting for replies (analogous to response feedback) from the space ship (analogous to an effector muscle).

Those of you who have read widely in psychology, or taken a number of courses, may especially enjoy Lashley's comment about his behaviorist colleague who remarks to him that "he had reached a stage where he could arise before an audience, turn his mouth loose, and go to sleep." This is a rather profound explanatory concept in and of itself.

Lashley considers the proposition that thinking is strictly a string of words, a weak form of the point of view expressed in the previous article by Watson. Lashley makes the basic point that we often know exactly what we wish to say without being able to find the right word. Surely this idea argues strongly against the notion that the thought *is* the word.

Lashley speaks on several occasions of "reverberatory circuits." The following analogy may help those of you not familiar with this concept. Suppose that a group of men stood in a circle all facing the center and each agreed to stand quietly unless someone tapped him on his right shoulder. When this occurred, he would turn to the left and tap his neighbor on the right shoulder. Our circle of men would stand quietly until an outsider tapped one of them on the right shoulder. He would then tap his neighbor, who would tap his neighbor, who would tap his neighbor, . . . , and a series of taps would start around the circle which would continue forever. This circle of men is quite closely analogous to a reverberatory circuit of neurons in which each neuron fires off the neuron after it which fires off the neuron after *it*, etc., with the impulse traveling round and round the circuit of neurons indefinitely.

As Lashley mentions, this article creates more problems than it solves; but, it undoubtedly moves us closer to a proper understanding of the marvelous workings of the human brain. Enjoy it.

THE PROBLEM OF SERIAL ORDER IN BEHAVIOR

K. S. Lashley

The previous speakers have approached our common problem by considering the properties of the elementary units of which we believe the cerebral structure to be built up. They have considered the kinds of neural integration or behavior which can be anticipated from those properties. The remaining members of the symposium have in their research been concerned chiefly with the analysis of complex behavior, seeking to derive general principles of neural integration from the infinitely complex products of that integration. Our common meeting ground is the faith to which we all subscribe, I believe, that the phenomena of behavior and of mind are ultimately describable in the concepts of the mathematical and physical sciences. In my discussion here, I have deliberately turned to the opposite extreme from the neuron and have chosen as a topic, one aspect of the most complex type of behavior that I know; the logical and orderly arrangement of thought and action. Our discussion so far has dealt

chiefly with the conditions of input and of immediate switching in the nervous mechanism, without explicit consideration of what is already going on within the system.

My principal thesis today will be that the input is never into a quiescent or static system, but always into a system which is already actively excited and organized. In the intact organism, behavior is the result of interaction of this background of excitation with input from any designated stimulus. Only when we can state the general characteristics of this background of excitation, can we understand the effects of a given input.

The unpronounceable Cree Indian word "kekawewechetushekamikowan-owow" is analyzed by Chamberlain (7) into the verbal root, *tusheka*, "to remain," and the various particles which modify it as follows: *ke(la)wow*, the first and last syllables, indicating second person plural; *ka*, a prefix of the future tense; *we*, a sort of imperative mode expressing a wish; *weche*, indicating conjunction of subject and object; *mik*, a suffix bringing the verb into agreement with a third person subject and second person object; and *owan*, a suffix indicating that the sub-

ject is inanimate and the object animate. A literal translation: "You will I wish together remain he-you it-man you" or, freely, "may it remain with you." This difference in structure between Cree and English illustrates an outstanding characteristic of verbal behavior; the occurrence of predetermined, orderly sequences of action which are unique for each language. In English the adjective precedes, in French it follows the noun which it modifies. In English the movement or action of the subject is expressed as early as possible after the subject; in German the expression of action may be postponed until all qualifying thoughts have been expressed. In a sentence discussing this subject, Pick (20) introduces fifty-five words between the subject and the principal verb. Each Chinese word, and to a lesser extent, each English word, stands as an unchanging unit. In the highly inflective languages, such as Sioux, the form of almost every word in the sentence may be altered, according to some attribute of the subject, as when two objects rather than one or several are discussed.

The study of comparative grammar is not the most direct approach to the physiology of the cerebral cortex, yet Fournié (10) has written, "Speech is the only window through which the physiologist can view the cerebral life." Certainly language presents in a most striking form the integrative functions that are characteristic of the cerebral cortex and that reach their highest development in human thought processes. Temporal integration is not found exclusively in language; the coordination of leg movements in insects, the song of birds, the control of trotting and pacing in a gaited horse, the rat running the maze, the architect designing a house, and

the carpenter sawing a board present a problem of sequences of action which cannot be explained in terms of successions of external stimuli.

Associative Chain Theories

In spite of the ubiquity of the problem, there have been almost no attempts to develop physiological theories to meet it. In fact, except among a relatively small group of students of aphasia, who have had to face questions of agrammatism, the problem has been largely ignored. It is not even mentioned in recent textbooks on neurophysiology or physiological psychology, nor is there any significant body of experimental studies bearing upon the problem. The spinal animal scarcely exhibits serial activity, so the physiologist may be excused for overlooking the phenomenon. On the other hand, psychologists have been concerned chiefly with the question of whether or not the organizing processes displayed in serial action are conscious, and very little with the organization itself. I have chosen to discuss the problem of temporal integration here, not with the expectation of offering a satisfactory physiological theory to account for it, but because it seems to me to be both the most important and also the most neglected problem of cerebral physiology. Temporally integrated actions do occur even among insects, but they do not reach any degree of complexity until the appearance of the cerebral cortex. They are especially characteristic of human behavior and contribute as much as does any single factor to the superiority of man's intelligence. A clearer formulation of the physiological problems which they raise should be of value, even though a solution of the problems is not yet in sight.

I shall consider first some of the questions raised by the structure of language, then turn to other forms of serial action for indications of the nature of the nervous mechanisms involved.

To the best of my knowledge, the only strictly physiological theory that has been explicitly formulated to account for temporal integration is that which postulates chains of reflexes, in which the performance of each element of the series provides excitation of the next. This conception underlays the "motor theories" of thinking which were advocated by several psychologists early in this century. Watson (26) sought to identify thought with inaudible movements of the vocal organs, linked together in associative chains. The peripheral chain theory of language was developed in greatest detail by Washburn (25). She distinguished what she called "successive movement systems" and, although she drew her examples from memorized series of nonsense syllables, her implication was that such series are typical of all language behavior. She defined a movement system as "a combination of movements so linked together that the stimulus furnished by the actual performance of certain movements is required to bring about other movements." She described speech as a succession of vocal acts in which the kinesthetic impulses from each movement serve as a unique stimulus for the next in the series (25, pages 11 ff.). Attempts to confirm these peripheral theories by mechanical (Thorsen, 23) or electrical (Max, 19) recording of muscular tensions have given no valid evidence in support of them. It should be noted that, at the time when the theories were proposed, it was generally believed that conduction in the nervous system is always

downstream from sense organ to muscle, and that muscular contraction must always follow promptly on stimulation. The existence of reverberatory circuits which could maintain central activity was scarcely suspected.

The introspective psychology which objected to such peripheral theories did not explicitly formulate an alternative neurological theory, but there is implicit in it a view that verbal thought is a simple chain of central processes in which each element serves to arouse the next by direct association. Titchener, for example, maintained that the meaning of a word (or of an auditory image in his system) consists of the chain of associations which it arouses; that it has no meaning until such a sequence has occurred. From this it must be inferred that he was thinking in terms of a simple associative chain, since no other relating process is suggested.

Objections to the Associative Chain Theory

A consideration of the structure of the sentence and of other motor sequences will show, I believe, that such interpretations of temporal organization are untenable and that there are, behind the overtly expressed sequences, a multiplicity of integrative processes which can only be inferred from the final results of their activity. There is an extensive controversial literature dealing with this inferred integrative activity. Pick (20) devotes almost his entire book, *Die agrammatischen Sprachstörungen*, to reviewing discussions of the subject. Most of this literature deals with the question of whether or not the integrative processes are conscious. Much of this is irrelevant to the present topic, but the advocates of so-called imageless

thought did present a great deal of material indicative of the complexity of the problem of thought structure. From this, and other evidence which I shall present, I believe that the production of speech involves the interaction of at least three, possibly four, major neurological systems which are interrelated but somewhat independently variable.

Let us start the analysis of the process with the enunciation of the word. Pronunciation of the word "right" consists first of retraction and elevation of the tongue, expiration of air and activation of the vocal cords; second, depression of the tongue and jaw; third, elevation of the tongue to touch the dental ridge, stopping of vocalization, and forceful expiration of air with depression of the tongue and jaw. These movements have no intrinsic order of association. Pronunciation of the word "tire" involves the same motor elements in reverse order. Such movements occur in all permutations. The order must therefore be imposed upon the motor elements by some organization other than direct associative connections between them. So, for the individual movements in writing or typing the word, finger strokes occur in all sorts of combinations. No single letter invariably follows *g*, and whether *gh*, *ga*, or *gu* is written depends upon a set for a larger unit of action, the word.

Words stand in relation to the sentence as letters do to the word; the words themselves have no intrinsic temporal "valence." The word "right," for example, is noun, adjective, adverb, and verb, and has four spellings and at least ten meanings. In such a sentence as "The mill-wright on my right thinks it right that some conventional rite should symbolize the right of every man to write as he

pleases," word arrangement is obviously not due to any direct associations of the word "right" itself with other words, but to meanings which are determined by some broader relations.

It has been found in studies of memorization of nonsense syllables that each syllable in the series has associations, not only with adjacent words in the series, but also with more remote words. The words in the sentence have, of course, associations with more remote words as well as with adjacent ones. However, the combination of such direct associations will not account for grammatical structure. The different positions of the word "right" in the illustrative sentence are determined by the meanings which the positions in relation to other words denote, but those meanings are given by other associations than those with the words in the spoken sentence. The word can take its position only when the particular one of its ten meanings becomes dominant. This dominance is not inherent in the words themselves.

From such considerations, it is certain that any theory of grammatical form which ascribes it to direct associative linkage of the words of the sentence overlooks the essential structure of speech. The individual items of the temporal series do not in themselves have a temporal "valence" in their associative connections with other elements. The order is imposed by some other agent.

This is true not only of language, but of all skilled movements or successions of movement. In the gaits of a horse, trotting, pacing, and single footing involve essentially the same pattern of muscular contraction in the individual legs. The gait is imposed by some mechanism in addition to the

direct relations of reciprocal innervation among the sensory-motor centers of the legs. The order in which the fingers of the musician fall on the keys or fingerboard is determined by the signature of the composition; this gives a *set* which is not inherent in the association of the individual movements.

The Determining Tendency

What then determines the order? The answer which seems most in accord with common sense is that the intention to act or the idea to be expressed determines the sequence. There are, however, serious difficulties for this solution. There is not much agreement among psychologists concerning the nature of the idea. The structuralist school, under the leadership of Titchener, held that the idea consists of mental images, often the auditory images of words, and the meanings are nothing but sequences of such images. Describing the role of images in his lecturing, Titchener wrote (24), "When there is any difficulty in exposition, a point to be argued *pro* and *con* or a conclusion to be brought out from the convergence of several lines of proof, I hear my own voice speaking just ahead of me." What solution of the lecture problem for the lazy man! He need not think but only listen to his own inner voice; to the chain of associated auditory images. A behaviorist colleague once remarked to me that he had reached a stage where he could arise before an audience, turn his mouth loose, and go to sleep. He believed in the peripheral chain theory of language. (This clearly demonstrates the superiority of behavioristic over introspective psychology. The behaviorist does not even have to listen to his own inner voice.)

Seriously, such positions offer no solution for the problem of temporal integration. Titchener finds his grammar ready made and does not even raise the question of the origin of the succession of images. The chain-reflex theory, while definite, is untenable.

The third view of the nature of the idea was developed by a group known as the "Würzburg School" (see Boring, 4); exponents of imageless thought. It held that some organization precedes any expression that can be discovered by introspective or objective means. Thought is neither muscular contraction nor image, but can only be inferred as a "determining tendency." At most, it is discovered as a vague feeling of pregnancy, of being about to have an idea, a *Bewusstseinslage*. It is not identical with the words which are spoken, for quite frequently no word can be recalled which satisfactorily expresses the thought, and we search a dictionary of synonyms until a word or phrase is found which does seem appropriate.

In his discussion of the relation of thought to speech, Pick (20) accepts this point of view, but he asserts further that the set or the idea does not have a temporal order; that all of its elements are cotemporal. Evidence in support of this conclusion comes, for example, from translation of one language into another which has a different sentence structure. I read a German sentence, pronouncing the German words with no thought of their English equivalents. I then give a free translation in English, without remembering a single word of the German text. Somewhere between the reading and free translation, the German sentence is condensed, the word order reversed, and expanded again into the different temporal order of English. According to Epstein (9), the polyglot shifts readily from one lan-

guage to another, expressing the same thought in either, without literal translation. The readiness with which the form of expression of an idea can be changed, the facility with which different word orders may be utilized to express the same thought, thus is further evidence that the temporal integration is not inherent in the preliminary organization of the idea.

The Schema of Order

The remaining alternative is that the mechanism which determines the serial activation of the motor units is relatively independent, both of the motor units and of the thought structure. Supporting evidence for this may be found in the mistakes of order, the slips and interferences which occur in writing and speaking. For some time I have kept records of errors in typing. A frequent error is the misplacing or the doubling of a letter. *These* is typed t-h-s-e-s, *look* as l-o-k-k, *ill* as i-i-l. Sometimes the set to repeat may be displaced by several words. The order is dissociated from the idea. Earlier, in preparing this paper, I wrote the phrase, "maintain central activities." I typed *min*, omitting the *a*, canceled this out and started again; *ama*. The impulse to insert the *a* now dominated the order. I struck out the *a* and completed the phrase, only to find that I had now also dropped the *a* from *activities*. This example suggests something of the complexity of the forces which are at play in the determination of serial order and the way in which conflicting impulses may distort the order, although the primary determining tendency, the idea, remains the same.

The polyglot, who has become proficient in a secondary language, who thinks in it and even dreams in it, may still tend to use the grammatical

structure of his native tongue. If, as in French, that tongue applies gender to inanimate things, the English pronouns referring to them may take the gender of the French equivalents, though the French nouns are not thought. The German postponement of the verb or the Magyar use of the past infinitive may be incorporated in the new language. In such cases, the structuring seems to be dissociated both from the content and from the simple associative connections of the words themselves.

The ease with which a new structure may be imposed on words is illustrated by the quickness with which children learn hog Latin. The form which I learned involved transposing the initial sound of each word to the end of the word and adding a long *a*. Thus—at-thay an-may oes-gay own-day e-thay eet-stray. Some children become very facile at such inversions of words, and re-structure new words without hesitation. From such considerations it seems to follow that syntax is not inherent in the words employed or in the idea to be expressed. It is a generalized pattern imposed upon the specific acts as they occur.

"Priming" of Expressive Units

There are indications that, prior to the internal or overt enunciation of the sentence, an aggregate of word units is partially activated or readied. Evidence for this comes also from "contaminations" of speech and writing. The most frequent typing errors are those of anticipation; the inclusion in the word being typed of some part of a word or word structure which should properly occur later in the sentence. It may be only a letter. Thus I wrote, *wrapid* writing, carrying the *w* from the second word to the first. Not infrequently words are

introduced which should occur much later in the sentence, often five or six words in advance.

In oral speech, Spoonerisms illustrate the same kind of contamination. The Spoonerism is most frequently an inversion of subject and object: "Let us always remember that waste makes haste." But it may be only a transposition of parts of the words: "Our queer old dean" for "our dear old queen." The frequency with which such contaminations occur is increased by haste, by distraction, by emotional tension, or by uncertainty and conflict as to the best form of expression. In some types of aphasia the tendency to disordered arrangement of words is greatly increased, and, in extreme cases, the attempt to speak results in a word hash with complete loss of grammatical organization. Professor Spooner, after whom such slips are named, was probably suffering from a mild form of aphasia. In these contaminations, it is as if the aggregate of words were in a state of partial excitation, held in check by the requirements of grammatical structure, but ready to activate the final common path, if the effectiveness of this check is in any way interfered with.

In his *Psychopathology of Everyday Life*, Freud has given numerous examples of similar contaminations of action outside the sphere of language. We do not need to accept his theories of censorship and suppression to account for such slips. They are of the same order as misplacements in typing and represent contaminations of co-existing, determining tendencies to action.

Such contaminations might be ascribed to differences in the relative strength of associative bonds between the elements of the act, and thus not evidence for pre-excitation of the ele-

ments or for simultaneous pre-excitation. However, the understanding of speech involves essentially the same problems as the production of speech and definitely demands the postulation of an after-effect or after-discharge of the sensory components for a significant time following stimulation. Thus, in the spoken sentence, "Rapid righting with his uninjured hand saved from loss the contents of the capsized canoe," the associations which give meaning to righting are not activated for at least 3 to 5 seconds after hearing the word.* I shall refer later to other evidence for such long after-discharge of sensory excitations. The fact of continued activation or after-discharge of receptive elements and their integration during this activation justifies the assumption of a similar process during motor organization. The processes of comprehension and production of speech have too much in common to depend on wholly different mechanisms.

Internal and Overt Speech

One other point with respect to the organization of speech: The earlier literature on aphasia emphasized the distinction of internal and overt speech. The aphemia of Broca and the pure motor aphasia of Wernicke and later writers were held to be a loss of the ability to enunciate without loss of ability to think in words and without paralysis of the organs of speech. The brain insult was assumed to affect only the transition from the thought to the enunciation

* Dr. Lashley ingeniously laid the groundwork for this three paragraphs earlier, when he mentions "wrapid writing." The audience all heard, "Rapid writing with his uninjured hand," etc. "Capsized canoe" required a complete and amusing about face. L. A. Jeffress.

of the word. We may doubt the existence of instances of such "pure" defects and question the reliability of the early clinical examinations in view of the more careful analyses that have been made since 1917, but the distinction of internal and overt speech is still valid and the transition still unexplained. Watson interpreted internal speech as inaudible movements of the vocal organs, and Jacobsen (15) and Max (19) have given evidence of changes in muscular tonus during verbal thinking or thought of movement. This is far from proving that the motor discharge is essential for the internal formation of words, however.

I once devised an instrument to record small movements of the tongue. Within the limits of its sensitivity, it showed that in silent thinking the tongue usually drops to the back of the mouth and shows no detectable movement. Verbal problems, such as the correct squaring of three-place numbers, could be carried out with no trace of overt movement. If, however, I urged the subject to hurry or if I slapped his face, his tongue came forward and showed movements corresponding to the syllabification of internal speech or of the computation he was performing. This I interpret as indicating that internal speech may be carried out wholly by processes within the nervous system, with some unessential discharge upon the final common path for vocal movements. Facilitation of the motor path, either by increased emotional tension or by "voluntary" reinforcement, increases its excitability until the same central circuits whose activity constitutes internal speech are able to excite the overt movements. This aspect of the language function is irrelevant to the problem of syntax or serial order, but

is important as illustrating a further point in the dynamics of the cerebrum. Many activities seem to require for their performance both a specific patterning and also a general facilitation, a rise in dynamic level. There are, I think, indications that hemiplegia and motor aphasia are primarily expressions of a low level of facilitation rather than a loss of specific integrative connections which are involved in the use of language or in the patterning of our movements. A monkey, for example, after ablation of the precentral gyrus may seem unable to use the arm at all, but if emotional excitement is raised above a certain level, the arm is freely used. As soon as the excitement dies down, the arm is again hemiplegic. I have seen something of the same sort in a human hemiplegic. The problem of the availability of memories, which was raised earlier in the discussion here, may find a partial solution in such fluctuations in dynamic level. In many of the organic amnesias the pattern of integration seems to be retained but can be reactivated only by an abnormally intense sensory or central reinforcement.

Generality of the Problem of Syntax

I have devoted so much time to discussion of the problem of syntax, not only because language is one of the most important products of human cerebral action, but also because the problems raised by the organization of language seem to me to be characteristic of almost all other cerebral activity. There is a series of hierarchies of organization: the order of vocal movements in pronouncing the word, the order of words in the sentence, the order of sentences in the para-

graph, the rational order of paragraphs in a discourse. Not only speech, but all skilled acts seem to involve the same problems of serial ordering, even down to the temporal coordination of muscular contractions in such a movement as reaching and grasping. Analysis of the nervous mechanisms underlying order in the more primitive acts may contribute ultimately to the solution even of the physiology of logic.

It is possible to designate, that is, to point to specific examples of, the phenomena of the syntax of movement that require explanation, although those phenomena cannot be clearly defined. A real definition would be a long step toward solution of the problem. There are at least three sets of events to be accounted for. First, the activation of the expressive elements (the individual words or adaptive acts) which do not contain the temporal relations. Second, the determining tendency, the set, or idea. This masquerades under many names in contemporary psychology, but is, in every case, an inference from the restriction of behavior within definite limits. Third, the syntax of the act, which can be described as an habitual order or mode of relating the expressive elements; a generalized pattern or schema of integration which may be imposed upon a wide range and a wide variety of specific acts. This is the essential problem of serial order; the existence of generalized schemata of action which determine the sequence of specific acts, acts which in themselves or in their associations seem to have no temporal valence.

I shall turn now to other phenomena of movement which may be more readily phrased in physiological terms and which may suggest some of the mechanisms underlying serial order.

Duration and Intensity of Nervous Discharge

A consideration of the control of extent and rate of movement supports the view that sensory factors play a minor part in regulating the intensity and duration of nervous discharge; that a series of movements is not a chain of sensory-motor reactions. The theory of control of movement which was dominant at the turn of the century assumed that, after a movement is initiated, it is continued until stopped by sensations of movement and position, which indicate that the limb has reached the desired position. This theory was opposed by a good bit of indirect evidence, such as that accuracy of movement is increased rather than diminished with speed. I had opportunity to study a patient who had a complete anesthesia for movements of the knee joint, as a result of a gunshot wound of the cord (16). In spite of the anesthesia, he was able to control the extent and speed of movements of flexion and extension of the knee quite as accurately as can a normal person.

The performance of very quick movements also indicates their independence of current control. "Whip-snapping" movements of the hand can be regulated accurately in extent, yet the entire movement, from initiation to completion, requires less than the reaction time for tactile or kinesthetic stimulation of the arm, which is about one-eighth of a second, even when no discrimination is involved. Such facts force the conclusion that an effector mechanism can be pre-set or primed to discharge at a given intensity or for a given duration, in independence of any sensory controls.

Central Control of Motor Patterns

This independence of sensory controls is true not only of intensity and duration of contraction of a synergic muscle group but is true also of the initiation and timing of contraction of the different muscles in a complex movement. The hand may describe a circular movement involving coordinated contractions of the muscles of the shoulder, elbow, and wrist in about $\frac{1}{10}$ second, and the stopping of movement at a given position, of course, is only a small fraction of that time. The finger strokes of a musician may reach sixteen per second in passages which call for a definite and changing order of successive finger movements. The succession of movements is too quick even for visual reaction time. In rapid sight reading it is impossible to read the individual notes of an arpeggio. The notes must be seen in groups, and it is actually easier to read chords seen simultaneously and to translate them into temporal sequence than to read successive notes in the arpeggio as usually written.

Sensory control of movement seems to be ruled out in such acts. They require the postulation of some central nervous mechanism which fires with predetermined intensity and duration or activates different muscles in predetermined order. . . .

Rhythmic Action

The simplest of the timing mechanisms are those controlling rhythmic activity. T. Graham Brown (5) first showed by his studies of de-afferented preparations that the rhythmic movements of respiration and progression are independent of peripheral stimu-

lation and are maintained by a central nervous mechanism of reciprocal innervation. He suggested that this mechanism of reciprocal innervation, rather than the simple reflex, is the unit of organization of the whole nervous system. He thus foreshadowed, in a way, the conception of reverberatory circuits which is coming to play so large a part in neurological theory today. Holst (13) has recently shown that the rhythmic movement of the dorsal fin of fishes is a compound of two superimposed rhythms, that of its own innervation and that of the pectoral fins. These two rhythms are centrally maintained.

Musical rhythms seem to be an elaboration of the same sort of thing. The time or beat is started and maintained at some definite rate, say 160 per minute. This rate is then imposed upon various activities. The fingers of the musician fall in multiples of the basic rate. If the leader of a quartet speeds up the time or retards, all the movements of the players change in rate accordingly. Not only the time of initiation but also the rate of movement is affected. The violinist, in a passage requiring the whole bow, will draw the bow from frog to tip at a uniform rate for the required number of beats, whether the tempo is fast or slow. With practiced violinists, the rate of movement is extremely accurate and comes out on the beat at the exact tip of the bow.

Superimposed on this primary rhythm is a secondary one of emphasis, giving the character of $\frac{3}{4}$, $\frac{4}{4}$, $\frac{6}{4}$, or other time. The mechanism of these rhythms can be simply conceived as the spread of excitation from some centers organized for reciprocal innervation, as a combination of the principles of Brown and of Holst. There are, however, still more com-

plicated rhythms in all music. That of the melodic line is most uniform. In much music, the melodic progression changes in 2, 4, or some multiple of 4 measures. In improvisation, the performer keeps no count of measures, yet comes out almost invariably in a resolution to the tonic of the key after some multiple of eight measures. Here a generalized pattern is impressed on the sequence, but it is a simpler pattern than that of grammatical structure. It only requires the recurrence of a pattern at certain rhythmic intervals; a pick-up of a specific pattern after so many timed intervals.

There are, in addition, still less regular rhythms of phrasing and emphasis. Parallels to these can be found in speech. The skilled extemporaneous speaker rounds his phrases and speaks with a definite though not regular rhythm.

The rhythms tend to spread to almost every other concurrent activity. One falls into step with a band, tends to breathe, and even to speak in time with the rhythm. The all pervasiveness of the rhythmic discharge is shown by the great difficulty of learning to maintain two rhythms at once, as in three against four with the two hands. The points to be emphasized here are the widespread effects of a rhythmic discharge indicating the involvement of almost the entire effector system, the concurrent action of different rhythmic systems, and the imposition of the rate upon both the initiation and speed of movement. Consideration of rhythmic activity and of spatial orientation forces the conclusion, I believe, that there exist in the nervous organization, elaborate systems of interrelated neurons capable of imposing certain types of integration upon a large number of widely spaced effector elements; in the one case trans-

mitting temporally spaced waves of facilitative excitation to all effector elements; in the other imparting a directional polarization to both receptor and effector elements. These systems are in constant action. They form a sort of substratum upon which other activity is built. They contribute to every perception and to every integrated movement.

Interaction of Temporal and Spatial Systems

Integration ascribed to the spatial distribution of excitations in the nervous system has been much more intensively studied than the temporal aspects of nervous activity. Theories of integration are based almost exclusively upon space properties, time entering only in theories of facilitation, inhibition, and after-discharge. In cerebral functions, however, it is difficult to distinguish between spatial and temporal functions. The eye is the only organ that gives simultaneous information concerning space in any detail. The shape of an object impressed on the skin can scarcely be detected from simultaneous pressure, but the same shape can readily be distinguished by touch when traced on the skin with a moving point or when explored by tactile scanning. The temporal sequence is readily translated into a spatial concept. Even for vision it might be questioned whether simultaneous stimulation gives rise directly to space concepts. The visual object is generally surveyed by eye movements, and its form is a reconstruction from such a series of excitations. Even with tachistoscopic exposures, the after-discharge permits a temporal survey, and, with visual fixation, shifts of attention provide an effective scanning.

Since memory traces are, we believe,

in large part static and persist simultaneously, it must be assumed that they are spatially differentiated. Nevertheless, reproductive memory appears almost invariably as a temporal sequence, either as a succession of words or of acts. Even descriptions of visual imagery (the supposed simultaneous reproductive memory in sensory terms) are generally descriptions of sequences, of temporal reconstructions from very fragmentary and questionable visual elements. Spatial and temporal order thus appear to be almost completely interchangeable in cerebral action. The translation from the spatial distribution of memory traces to temporal sequence seems to be a fundamental aspect of the problem of serial order.

I spoke earlier of the probability of a partial activation or priming of aggregates of words before the sentence is actually formulated from them. There is a great deal of evidence for such preliminary facilitation of patterns of action in studies of reaction time and of word association. Reaction time, in general, is reduced by preliminary warning or by instructions which allow the subject to prepare for the specific act required. In controlled association experiments, the subject is instructed to respond to the stimulus word by a word having a certain type of relation to it, such as the opposite or a part of which the stimulus is the whole; black-white, apple-seed. The result is an attitude or set which causes that particular category to dominate the associative reaction. Whether such preliminary reinforcement is to be ascribed to accumulation of excitatory state, as defined by Sherrington (21), or to some other physiological process, the facts of behavior assure that it is a genuine phenomenon and plays a de-

cisive role in determining the character of the response.

Once the existence of such states of partial activation is recognized, their possible role in temporal integration must be considered. There are indications that one neural system may be held in this state of partial excitation while it is scanned by another. Here is an example. A series of four to six numbers is heard: 3-7-2-9-4. This is within the attention or memory span and is almost certainly not remembered in the sense in which one's telephone number is remembered, for memory of it is immediately wiped out by a succeeding series of numbers. While it is retained in this unstable way, subject to retroactive inhibition, the order of the numbers can be reassembled: 3-7-2-9-4, 3-2-7-9-4, 4-9-2-7-3, and the like. It is as if, in this case, a rhythmic alternation can suppress alternate items, or a direction of arousal can be applied to the partially excited system. Another example which illustrates even more clearly the spatial characteristics of many memory traces is the method of comultiplication, used in rapid mental calculation. In attempts to play a melody backward, we have a further illustration. I find that I can do it only by visualizing the music spatially and then reading it backward. I cannot auditorily transform even "Yankee Doodle" into its inverse without some such process, but it is possible to get a spatial representation of the melody and then to scan the spatial representation. The scanning of a spatial arrangement seems definitely to determine, in such cases, the order of procedure. Two assumptions are implied by this. First, the assumption is that the memory traces are associated, not only with other memory traces, but also with the system of space coordi-

nates. By this I do not mean that the engram has a definite location in the brain; our experiments show conclusively that such is not the case. Rather, when the memory trace is formed it is integrated with directional characters of the space system, which give it position in reference to other associated traces. Second, the assumption is that these space characters of the memory trace can be scanned by some other level of the coordinating system and so transformed into succession.

This is as far as I have been able to go toward a theory of serial order in action. Obviously, it is inadequate. The assumptions concerning spatial representation and temporal representation may even beg the question, since no one can say whether spatial or temporal order is primary. Furthermore, such determining tendencies as the relation of attribute to object, which gives the order of adjective and noun, do not seem to be analyzable into any sort of spatial structure or for that matter, into any consistent relationship. I have tried a number of assumptions concerning the selective mechanism of grammatical form (spatial relations, the relative intensity or prominence of different words in the idea, and so on) but I have never been able to make an hypothesis which was consistent with any large number of sentence structures. Nevertheless, the indications which I have cited, that elements of the sentence are readied or partially activated before the order is imposed upon them in expression, suggest that some scanning mechanism must be at play in regulating their temporal sequence. The real problem, however, is the nature of the selective mechanism by which the particular acts are picked out in this scanning process, and to this problem I have no answer.

Such speculations concerning temporal and spatial systems do little more than illustrate a point of view concerning nervous organization which is, I believe, more consistent both with what is known of the histology and elementary physiology of the brain and also with behavior phenomena than are the more widely current theories of simple associative chains of reactions.

Nearly forty years ago Becher (2, page 243) wrote: "There is no physiological hypothesis which can explain the origin and relations of temporal forms in mental life; indeed, there is no hypothesis which even foreshadows the possibility of such an explanation." The situation is little better today, but I do feel that changing conceptions of the fundamental organization of the nervous system offer more hope for a solution of such problems than did the physiological knowledge available when Becher wrote. However, we are still very far from being able to form an explicit explanation of temporal structure.

The Fundamental Mechanism of Integration

Neurological theory has been dominated by the belief that the neurons of the central nervous system are in an inactive or resting state for the greater part of the time; that they are linked in relatively isolated conditioned reflex arcs and that they are activated only when the particular reactions for which they are specifically associated are called out. Such a view is incompatible both with the widespread effects of stimulation which can be demonstrated by changes in tonus and also with recent evidence from electrical recording of nervous activity. It is now practically certain

that all the cells of the cerebrospinal axis are being continually bombarded by nerve impulses from various sources and are firing regularly, probably even during sleep. The nervous activity which they in turn elicit depends upon the current physiological state of the neurons with which they are connected. It is probably not far from the truth to say that every nerve cell of the cerebral cortex is involved in thousands of different reactions. The cortex must be regarded as a great network of reverberatory circuits, constantly active. A new stimulus, reaching such a system, does not excite an isolated reflex path but must produce widespread changes in the pattern of excitation throughout a whole system of already interacting neurons.

The facts of cerebral structure support such a view. The cortex is composed chiefly of neurons with short axons. LeGros Clark (8) has found for the striate area of the monkey that Marchi degeneration extends for only a short distance from a point of injury. In the striate area of the rat, I have never been able to trace degeneration beyond three or four cell diameters from the margin of a lesion, and I believe that this lack of long transcortical fibers is true of other areas as well as of the visual cortex. Visual perception reveals close integration of different parts of the striate areas in spite of the absence of long association fibers. In the visual cortex of the rat there are only 19 neurons for each afferent fiber. To produce the animal's visual acuity, all of the afferent fibers must be firing continually. There are approximately 34,000 cell bodies in the lateral geniculate nucleus of the rat, and the minimum number of visual units necessary to produce the visual acuity of the rat is actually above this figure. (The acuity is determined by direct experi-

mental tests.) These figures should be of interest in relation to the numerical values cited by Dr. von Neumann. The number of cells in the visual cortex of the rat is only about 10^6 , and in some of my experiments where I have removed the greater part of the visual cortex the capacity for discrimination of visual forms has been retained when no more than 20,000 cells of the visual cortex remain. There is also evidence that no part of the cerebral cortex except the visual areas is essential for visual perception and memory.

Dr. Lorente de Nó: What is the number of afferents in the optic nerve?

Dr. Lashley: There are 290,000 afferents in the optic nerve of the rat, and the figure is reduced to 34,000 in the lateral geniculate. The actual numbers are 9,000,000 myoids, 290,000 ganglion cells, and 34,000 cells in the lateral geniculate. That may include cells with short axons also. There are about 125,000 cells in each of the five layers of the cortex. These figures are for one eye and hemisphere.

Dr. von Neumann: In the human being the corresponding number is about 125,000,000 for the first, isn't it?

Dr. Lashley: I know of no figure for that level.

Dr. Lorente de Nó: It hasn't been analyzed in any way.

Dr. von Neumann: The optic nerve corresponds to the second one?

Dr. Lashley: Yes. The axons of the ganglion cells pass through the optic nerves. There is an average of 300 visual cells firing into each central pathway. There are fewer than 5 cells in the receptive layer of the visual cortex of the rat for each afferent fiber of the optic radiation and only 19 cells per afferent fiber in the entire visual cortex. Since the visual acuity of the rat requires that all of the 34,000 cells of the radiation be firing constantly, it seems certain that all of the neurons within the striate areas, the visual cortex, must be firing constantly. There is a good bit of evidence that all of the integrative functions of vision are carried out within the striate areas. In the rat, I have removed from one or an-

other animal, practically every other part of the isocortex without disturbing visual perception or memory. With monkeys I have removed the supposed visual associative areas without producing any significant loss of visual functions.

These facts lead to the conclusion that the same cells in the visual cortex participate in a great variety of activities. Practically all of the cells of the area must be fired by every visual stimulation, and these same cells must be the ones which retain the visual memories. The conclusion follows that differential responses depend upon the pattern of cells which are excited in combination. The visual cortex is a network of cells of short axon without long interconnections between its parts or with other cortical areas. Its integrative functions are an expression of the properties of such a network.

The same conception must be applied to other cortical areas. There are, of course, long association tracts in the cortex, such as the corpus callosum, the superior longitudinal fasciculus, and the temporo-frontal tracts. Once, 26 years ago, I suggested facetiously that these might be only skeletal structures, since I could find no function for them. No important functions of these tracts have yet been demonstrated. Section of the corpus callosum produces only a slight slowing of reaction time, ipsilateral as well as contralateral (Akelaitis, 1); section of occipito-frontal fibers produces, perhaps, a temporary disturbance of visual attention but no other symptoms.

The integrative functions seem to be carried out as well without as with the main associative tracts. The major integrative functions must, therefore, be carried out by the network of cells of short axon. The properties of such networks of cells must be analyzed before the mechanisms of the cerebral cortex can be understood. Something can be inferred from the characteristics of excitability of cells and their arrangement in recurrent loops. If, as seems a necessary conclusion from the histology of the striate area, all of the cells of the network are subject to constant excitation and are firing whenever they recover from the refractory state, then mutual interference of circuits will produce complicated patterns throughout the area, patterns which will stabilize in the absence of differential stimulation, as is perhaps indicated by the regularity of the alpha rhythm. Any new afferent impulses reaching the area can only produce a reorganization of the existing pattern. What happens at any particular point in the system, as at an efferent neuron, is the statistical outcome of the interaction of myriads of neurons, not of the transmission of impulses over a restricted path, of which that efferent cell forms a link. It is possible to isolate parts of the system by operative means or by anesthetics and so to get a one-to-one relation of stimulus locus and responding muscles, from which the reflex mechanism has been inferred. As Goldstein (12) has pointed out, however, the parts isolated in the reflex are influenced by a multiplicity of effects in the intact organism of which there is little or no trace in the isolated preparation.

I can best illustrate this conception of nervous action by picturing the brain as the surface of a lake. The prevailing breeze carries small ripples

in its direction, the basic polarity of the system. Varying gusts set up crossing systems of waves, which do not destroy the first ripples, but modify their form, a second level in the system of space coordinates. A tossing log with its own period of submersion sends out periodic bursts of ripples, a temporal rhythm. The bow wave of a speeding boat momentarily sweeps over the surface, seems to obliterate the smaller waves yet leaves them unchanged by its passing, the transient effect of a strong stimulus. Wave motion is not an adequate analogy because the medium which conveys the waves is uniform, whereas the nerve cells have their individual characteristics of transmission which at every point may alter the character of the transmitted pattern.

The great number of axon terminations on every nerve cell has not been considered in theories of integration. It implies, of course, that the cell can be fired by impulses from a variety of sources. But it also suggests another possibility, more fruitful for understanding of integrative processes. A nerve impulse arriving over a single axon terminal may not fire the cell but may modify its excitability to impulses from other sources. In an elaborate system of neurons such subthreshold effects might establish a pattern of facilitation which would determine the combination of cells fired by subsequent excitations. The space coordinate system and various types of *set* or priming may be pictured as patterns of subthreshold facilitation pervading the network of neurons which is activated by the more specific external stimulus.

Such a view of the mechanism of nervous action certainly does not simplify the problems nor does it as yet provide any clue to the structuring

that constitutes the set or determining tendency, or to the nature of such relations as are implied in the attribute-object, opposites, or other abstract concepts. A few relations seem reducible to spatial terms, part-whole, for example, but even for these there is no clear conception of the neural basis of their space properties. These considerations do not, I believe, contradict fundamentally the basic conceptions that have been formulated by Dr. McCulloch. They do, however, indicate a direction of necessary elaboration. The nets active in rhythmic and spatial organization are apparently almost coextensive with the nervous system. The analysis must be extended to the properties of such nets; the way in which they are broken up into reactive patterns in the spread of excitation, to give, for example, directional propagation or its equivalent. I strongly suspect that many phenomena of generalization, both sensory and conceptual, are products, not of simple switching, but of interaction of complex patterns of organization within such systems.

Summary

The problems of the syntax of action are far removed from anything which we can study by direct physiological methods today, yet in attempting to formulate a physiology of the cerebral cortex we cannot ignore them. Serial order is typical of the problems raised by cerebral activity; few, if any, of the problems are simpler or promise easier solution. We can, perhaps, postpone the fatal day when we must face them, by saying that they are too complex for present analysis, but there is danger here of constructing a false picture of those processes that we believe to be simpler.

I am coming more and more to the conviction that the rudiments of every human behavioral mechanism will be found far down in the evolutionary scale and also represented even in primitive activities of the nervous system. If there exist, in human cerebral action, processes which seem fundamentally different or inexplicable in terms of our present construct of the elementary physiology of integration, then it is probable that that construct is incomplete or mistaken, even for the levels of behavior to which it is applied.

In spite of its present inadequacy, I feel that the point of view which I have sketched here holds some promise of a better understanding of cerebral integration. Attempts to express cerebral function in terms of the concepts of the reflex arc, or of associated chains of neurons, seem to me doomed to failure because they start with the assumption of a static nervous system. Every bit of evidence available indicates a dynamic, constantly active system, or, rather, a composite of many interacting systems, which I have tried to illustrate at a primitive level by rhythm and the space coordinates. Only when methods of analysis of such systems have been devised will there be progress toward understanding of the physiology of the cerebral cortex.

References

1. AKELAITIS, A. J., Studies on the corpus callosum. ii. The higher visual functions in each homonymous field following complete section of the corpus callosum. *Arch. Neurol. Psychiat.*, 1941, **45**, 788-796.
2. BECHER, E., *Gehirn und Seele*. Heidelberg, 1911.

3. BETHE, A., Plastizität und Zentrenlehre. *Handb. d. norm. u. path. Physiol.*, 1931, 15 (zweite H.), 1175-1220.
4. BORING, E. G., *A history of experimental psychology*. New York: Century Co., 1929.
5. BROWN, T. G., On the nature of the fundamental activity of the nervous centers. *J. Physiol.*, 1941, 48, 18-46.
6. BUDDENBROCK, W. v., Die Rhythmus der Schreitbewegungen der Stabheuschrecke *Dyxippus*. *Biol. Centralb.*, 1921, 41, 41-48.
7. CHAMBERLAIN, A. F., Indians, North American. *Enc. Brit.*, 1911, 14, 452-482.
8. CLARK, W. E. LE GROS, Observations on the associative fiber system of the visual cortex and the central representation of the retina. *J. Anat. London*, 1941, 75, 225-236.
9. EPSTEIN, I., *La pensée et la polyglotie*. Paris: Payot et Cie (no date).
10. FOURNIÉ, *Essai de psychologie*. Paris, 1887.
11. FRITSCH, G., and HITZIG, E., Ueber die elektrische Erregbarkeit des Grosshirns. *Arch. f. Anat. u. Physiol.*, 1870, pp. 300-332.
12. GOLDSTEIN, K., *The organism*. Boston: Ginn & Co., 1939.
13. HOLST, N. v., Vom Wesen der Ordnung im Zentralnervensystem. *Die Naturwissenschaften*, 1937, 25, 625-631, 641-647.
14. INGEBRITZEN, O. C., Coordinating mechanisms of the spinal cord. *Genet. Psychol. Monogr.*, 1933, 13, 483-555.
15. JACOBSEN, E., Electrophysiology of mental activities. *Amer. J. Psychol.*, 1932, 44, 677-694.
16. LASHLEY, K. S., The accuracy of movement in the absence of excitation from the moving organ. *Amer. J. Physiol.*, 1917, 43, 169-194.
17. LASHLEY, K. S., The mechanism of vision. xvii. Autonomy of the visual cortex. *J. Genet. Psychol.*, 1942, 60, 197-221.
18. LASHLEY, K. S., The mechanism of vision. xviii. Effects of destroying the visual "associative areas" in the monkey. *Genet. Psychol. Monogr.*, 1948, 37, 107-166.
19. MAX, L. W., Experimental study of the motor theory of consciousness. IV. *J. Comp. Psychol.*, 1937, 24, 301-344.
20. PICK, A., *Die agrammatischen Sprachstörungen*. Berlin, 1913.
21. SHERRINGTON, C. S., *The integrative action of the nervous system*. London: Constable, 1906.
22. SHERRINGTON, C. S., Some functional problems attaching to convergence. *Proc. Roy Soc., B.*, 1929, 105, 332-362.
23. THORSEN, A. M., The relation of tongue movements to internal speech. *J. Exp. Psychol.*, 1925, 8, 1-32.
24. TITCHENER, E. B., *Lectures on the experimental psychology of the thought processes*. New York: The Macmillan Co., 1909.
25. WASHBURN, M. F., *Movement and mental imagery*. Boston: Houghton Mifflin, 1916.
26. WATSON, J. B., Is thinking merely the action of the language mechanisms? *Brit. J. Psychol.*, 1920, 11, 86-104.

INTRODUCTION

This article, by the late Keller Breland and his wife Marian, describes some of the pitfalls that they encountered when they left academic psychology to become animal trainers. *Before* reading this paper, you could profitably consider the following questions: Can any desired kind of simple behavior be trained by appropriate use of reward for that behavior? Is it probably the case that most learned behavior is acquired according to the simple laws of operant conditioning? Furthermore, is most behavior learned rather than innate?

The Brelands report that they began their commercial enterprise with the tacit assumption that "all responses are about equally conditionable to all stimuli." There is no doubt that this assumption has often been made, implicitly and explicitly, in the literature of psychology. In the following account you will find some reasons for doubting the generality of the assumption. In the article after this one, by Dobrzecka, *et al.*, you will find some further evidence, this time gathered in the laboratory, for questioning the general assumption. Articles of this sort can remind us that it would be dangerous to ignore altogether the possibilities of human instinctive behavior patterns. For example, when you read the selection by Azrin, in the *Frontiers* section, consider whether there is a human analogue of the pain-produced-aggression which Azrin has studied in animals. Those concerned about tyrannical control of populations by dictators in possession of the laws of behavior may be cheered by the Brelands' observation that "the animal simply does not do what he has been conditioned to do."

THE MISBEHAVIOR OF ORGANISMS

Keller Breland
Marian Breland

There seems to be a continuing realization by psychologists that perhaps the white rat cannot reveal everything there is to know about behavior. Among the voices raised on this topic, Beach (1950) has emphasized the necessity of widening the range of species subjected to experimental techniques and conditions. However, psychologists as a whole do not seem to be heeding these admonitions, as Whalen (1961) has pointed out.

Perhaps this reluctance is due in part to some dark precognition of what they might find in such investigations, for the ethologists Lorenz (1950, p. 233) and Tinbergen (1951, p. 6) have warned that if psychologists are to understand and predict the behavior of organisms, it is essential that they become thoroughly familiar with the instinctive behavior patterns of each new species they essay to study. Of course, the Watsonian or neobehavioristically oriented experimenter is apt to consider "instinct" an ugly

word. He tends to class it with Hebb's (1960) other "seditious notions" which were discarded in the behavioristic revolution, and he may have some premonition that he will encounter this *bête noire* in extending the range of species and situations studied.

We can assure him that his apprehensions are well grounded. In our attempt to extend a behavioristically oriented approach to the engineering control of animal behavior by operant conditioning techniques, we have fought a running battle with the seditious notion of instinct.* It might be of some interest to the psychologist to know how the battle is going and to learn something about the nature of the adversary he is likely to meet if and when he tackles new species in new learning situations.

Our first report (Breland & Breland, 1951) in the *American Psychologist*, concerning our experiences in controlling animal behavior, was wholly affirming and optimistic, saying in essence that the principles derived from the laboratory could be applied to the extensive control of behavior.

* In view of the fact that instinctive behaviors may be common to many zoological species, we consider *species specific* to be a sanitized misnomer, and prefer the possibly septic adjective *instinctive*.

We are grateful to Mrs. Breland and the publisher for their kind permission to include this selection. Reprinted from Keller and Marian Breland's "The Misbehavior of Organisms," American Psychologist, 16, 1961, pp. 681-84. Copyright © 1961 by the American Psychological Association and reproduced by permission.

ior under nonlaboratory conditions throughout a considerable segment of the phylogenetic scale.

When we began this work, it was our aim to see if the science would work beyond the laboratory, to determine if animal psychology could stand on its own feet as an engineering discipline. These aims have been realized. We have controlled a wide range of animal behavior and have made use of the great popular appeal of animals to make it an economically feasible project. Conditioned behavior has been exhibited at various municipal zoos and museums of natural history and has been used for department store displays, for fair and trade convention exhibits, for entertainment at tourist attractions, on television shows, and in the production of television commercials. Thirty-eight species, totaling over 6,000 individual animals, have been conditioned, and we have dared to tackle such unlikely subjects as reindeer, cockatoos, raccoons, porpoises, and whales.

Emboldened by this consistent reinforcement, we have ventured further and further from the security of the Skinner box. However, in this cavalier extrapolation, we have run afoul of a persistent pattern of discomforting failures. These failures, although disconcertingly frequent and seemingly diverse, fall into a very interesting pattern. They all represent breakdowns of conditioned operant behavior. From a great number of such experiences, we have selected, more or less at random, the following examples.

The first instance of our discomfort might be entitled, "What Makes Sammy Dance?" In the exhibit in which this occurred, the casual observer sees a grown bantam chicken emerge from a retaining compartment

when the door automatically opens. The chicken walks over about 3 feet, pulls a rubber loop on a small box which starts a repeated auditory stimulus pattern (a four-note tune). The chicken then steps up onto an 18-inch, slightly raised disc, thereby closing a timer switch, and scratches vigorously, round and round, over the disc for 15 seconds, at the rate of about two scratches per second until the automatic feeder fires in the retaining compartment. The chicken goes into the compartment to eat, thereby automatically shutting the door. The popular interpretation of this behavior pattern is that the chicken has turned on the "juke box" and "dances."

The development of this behavioral exhibit was wholly unplanned. In the attempt to create quite another type of demonstration which required a chicken simply to stand on a platform for 12–15 seconds, we found that over 50% developed a very strong and pronounced scratch pattern, which tended to increase in persistence as the time interval was lengthened. (Another 25% or so developed other behaviors—pecking at spots, etc.) However, we were able to change our plans so as to make use of the scratch pattern, and the result was the "dancing chicken" exhibit described above.

In this exhibit the only real contingency for reinforcement is that the chicken must depress the platform for 15 seconds. In the course of a performing day (about 3 hours for each chicken) a chicken may turn out over 10,000 unnecessary, virtually identical responses. Operant behaviorists would probably have little hesitancy in labeling this an example of Skinnerian "superstition" (Skinner, 1948) or "medicating" behavior, and we list it first to whet their explanatory appetite.

However, a second instance involving a raccoon does not fit so neatly into this paradigm. The response concerned the manipulation of money by the raccoon (who has "hands" rather similar to those of the primates). The contingency for reinforcement was picking up the coins and depositing them in a 5-inch metal box.

Raccoons condition readily, have good appetites, and this one was quite tame and an eager subject. We anticipated no trouble. Conditioning him to pick up the first coin was simple. We started out by reinforcing him for picking up a single coin. Then the metal container was introduced, with the requirement that he drop the coin into the container. Here we ran into the first bit of difficulty: he seemed to have a great deal of trouble letting go of the coin. He would rub it up against the inside of the container, pull it back out, and clutch it firmly for several seconds. However, he would finally turn it loose and receive his food reinforcement. Then the final contingency: we put him on a ratio of 2, requiring that he pick up both coins and put them in the container.

Now the raccoon really had problems (and so did we). Not only could he not let go of the coins, but he spent seconds, even minutes, rubbing them together (in a most miserly fashion), and dipping them into the container. He carried on this behavior to such an extent that the practical application we had in mind—a display featuring a raccoon putting money in a piggy bank—simply was not feasible. The rubbing behavior became worse and worse as time went on, in spite of nonreinforcement.

For the third instance, we return to the gallinaceous birds. The observer sees a hopper full of oval plastic

capsules which contain small toys, charms, and the like. When the S_D (a light) is presented to the chicken, she pulls a rubber loop which releases one of these capsules onto a slide about 16 inches long, inclined at about 30 degrees. The capsule rolls down the slide and comes to rest near the end. Here one or two sharp, straight pecks by the chicken will knock it forward off the slide and out to the observer, and the chicken is then reinforced by an automatic feeder. This is all very well—most chickens are able to master these contingencies in short order. The loop pulling presents no problems; she then has only to peck the capsule off the slide to get her reinforcement.

However, a good 20% of all chickens tried on this set of contingencies fail to make the grade. After they have pecked a few capsules off the slide, they begin to grab at the capsules and drag them backwards into the cage. Here they pound them up and down on the floor of the cage. Of course, this results in no reinforcement for the chicken, and yet some chickens will pull in over half of all the capsules presented to them.

Almost always this problem behavior does not appear until after the capsules begin to move down the slide. Conditioning is begun with stationary capsules placed by the experimenter. When the pecking behavior becomes strong enough, so that the chicken is knocking them off the slide and getting reinforced consistently, the loop pulling is conditioned to the light. The capsules then come rolling down the slide to the chicken. Here most chickens, who before did not have this tendency, will start grabbing and shaking.

The fourth incident also concerns a chicken. Here the observer sees a

chicken in a cage about 4 feet long which is placed alongside a miniature baseball field. The reason for the cage is the interesting part. At one end of the cage is an automatic electric feed hopper. At the other is an opening through which the chicken can reach and pull a loop on a bat. If she pulls the loop hard enough the bat (solenoid operated) will swing, knocking a small baseball up the playing field. If it gets past the miniature toy players on the field and hits the back fence, the chicken is automatically reinforced with food at the other end of the cage. If it does not go far enough, or hits one of the players, she tries again. This results in behavior on an irregular ratio. When the feeder sounds, she then runs down the length of the cage and eats.

Our problems began when we tried to remove the cage for photography. Chickens that had been well conditioned in this behavior became wildly excited when the ball started to move. They would jump up on the playing field, chase the ball all over the field, even knock it off on the floor and chase it around, pecking it in every direction, although they had never had access to the ball before. This behavior was so persistent and so disruptive, in spite of the fact that it was never reinforced, that we had to reinstate the cage.

The last instance we shall relate in detail is one of the most annoying and baffling for a good behaviorist. Here a pig was conditioned to pick up large wooden coins and deposit them in a large "piggy bank." The coins were placed several feet from the bank and the pig required to carry them to the bank and deposit them, usually four or five coins for one reinforcement. (Of course, we started out with one coin, near the bank.)

Pigs condition very rapidly, they have no trouble taking ratios, they have ravenous appetites (naturally), and in many ways are among the most tractable animals we have worked with. However, this particular problem behavior developed in pig after pig, usually after a period of weeks or months, getting worse every day. At first the pig would eagerly pick up one dollar, carry it to the bank, run back, get another, carry it rapidly and neatly, and so on, until the ratio was complete. Thereafter, over a period of weeks the behavior would become slower and slower. He might run over eagerly for each dollar, but on the way back, instead of carrying the dollar and depositing it simply and cleanly, he would repeatedly drop it, root it, drop it again, root it along the way, pick it up, toss it up in the air, drop it, root it some more, and so on.

We thought this behavior might simply be the dilly-dallying of an animal on a low drive. However, the behavior persisted and gained in strength in spite of a severely increased drive—he finally went through the ratios so slowly that he did not get enough to eat in the course of a day. Finally it would take the pig about 10 minutes to transport four coins a distance of about 6 feet. This problem behavior developed repeatedly in successive pigs.

There have also been other instances: hamsters that stopped working in a glass case after four or five reinforcements, porpoises and whales that swallow their manipulanda (balls and inner tubes), cats that will not leave the area of the feeder, rabbits that will not go to the feeder, the great difficulty in many species of conditioning vocalization with food reinforcement, problems in conditioning a kick in a cow, the failure to get

appreciably increased effort out of the ungulates with increased drive, and so on. These we shall not dwell on in detail, nor shall we discuss how they might be overcome.

These egregious failures came as a rather considerable shock to us, for there was nothing in our background in behaviorism to prepare us for such gross inabilities to predict and control the behavior of animals with which we had been working for years.

The examples listed we feel represent a clear and utter failure of conditioning theory. They are far from what one would normally expect on the basis of the theory alone. Furthermore, they are definite, observable; the diagnosis of theory failure does not depend on subtle statistical interpretations or on semantic legerdemain—the animal simply does not do what he has been conditioned to do.

It seems perfectly clear that, with the possible exception of the dancing chicken, which could conceivably, as we have said, be explained in terms of Skinner's superstition paradigm, the other instances do not fit the behavioristic way of thinking. Here we have animals, after having been conditioned to a specific learned response, gradually drifting into behaviors that are entirely different from those which were conditioned. Moreover, it can easily be seen that these particular behaviors to which the animals drift are clear-cut examples of instinctive behaviors having to do with the natural food getting behaviors of the particular species.

The dancing chicken is exhibiting the gallinaceous birds' scratch pattern that in nature often precedes ingestion. The chicken that hammers capsules is obviously exhibiting instinctive behavior having to do with breaking open of seed pods or the

killing of insects, grubs, etc. The raccoon is demonstrating so-called "washing behavior." The rubbing and washing response may result, for example, in the removal of the exoskeleton of a crayfish. The pig is rooting or shaking—behaviors which are strongly built into this species and are connected with the food getting repertoire.

These patterns to which the animals drift require greater physical output and therefore are a violation of the so-called "law of least effort." And most damaging of all, they stretch out the time required for reinforcement when nothing in the experimental setup requires them to do so. They have only to do the little tidbit of behavior to which they were conditioned—for example, pick up the coin and put it in the container—to get reinforced immediately. Instead, they drag the process out for a matter of minutes when there is nothing in the contingency which forces them to do this. Moreover, increasing the drive merely intensifies this effect.

It seems obvious that these animals are trapped by strong instinctive behaviors, and clearly we have here a demonstration of the prepotency of such behavior patterns over those which have been conditioned.

We have termed this phenomenon "instinctive drift." The general principle seems to be that wherever an animal has strong instinctive behaviors in the area of the conditioned response, after continued running the organism will drift toward the instinctive behavior to the detriment of the conditioned behavior and even to the delay or preclusion of the reinforcement. In a very boiled-down, simplified form, it might be stated as "learned behavior drifts toward instinctive behavior."

All this, of course, is not to disparage the use of conditioning techniques, but is intended as a demonstration that there are definite weaknesses in the philosophy underlying these techniques. The pointing out of such weaknesses should make possible a worthwhile revision in behavior theory.

The notion of instinct has now become one of our basic concepts in an effort to make sense of the welter of observations which confront us. When behaviorism tossed out instinct, it is our feeling that some of its power of prediction and control were lost with it. From the foregoing examples, it appears that although it was easy to banish the Instinctivists from the science during the Behavioristic Revolution, it was not possible to banish instinct so easily.

And if, as Hebb suggests, it is advisable to reconsider those things that behaviorism explicitly threw out, perhaps it might likewise be advisable to examine what they tacitly brought it—the hidden assumptions which led most disastrously to these breakdowns in the theory.

Three of the most important of these tacit assumptions seem to us to be: that the animal comes to the laboratory as a virtual *tabula rasa*, that species differences are insignificant, and that all responses are about equally conditionable to all stimuli.

It is obvious, we feel, from the foregoing account, that these assumptions are no longer tenable. After 14 years of continuous conditioning and observation of thousands of animals, it is our reluctant conclusion that the behavior of any species cannot be adequately understood, predicted, or controlled without knowledge of its instinctive patterns, evolutionary history, and ecological niche.

In spite of our early successes with the application of behavioristically oriented conditioning theory, we readily admit now that ethological facts and attitudes in recent years have done more to advance our practical control of animal behavior than recent reports from American "learning labs."

Moreover, as we have recently discovered, if one begins with evolution and instinct as the basic format for the science, a very illuminating viewpoint can be developed which leads naturally to a drastically revised and simplified conceptual framework of startling explanatory power (to be reported elsewhere).

It is hoped that this playback on the theory will be behavioral technology's partial repayment to the academic science whose impeccable empiricism we have used so extensively.

References

- BEACH, F. A., The snark was a boojum. *Amer. Psychologist*, 1950, 5, 115-124.
- BRELAND, K., and M. BRELAND, A field of applied animal psychology. *Amer. Psychologist*, 1951, 6, 202-204.
- HEBB, D. O., The American revolution. *Amer. Psychologist*, 1960, 15, 735-745.
- LORENZ, K., Innate behaviour patterns. In: *Symposia of the Society for Experimental Biology*. No. 4. *Physiological mechanisms in animal behaviour*. New York: Academic Press, 1950.
- SKINNER, B. F., Superstition in the pigeon. *J. exp. Psychol.*, 1948, 38, 168-172.
- TINBERGEN, N., *The study of instinct*. Oxford: Clarendon, 1951.
- WHALEN, R. E., Comparative psychology. *Amer. Psychologist*, 1961, 16, 84.

INTRODUCTION

In this short selection the authors present some data very relevant to the question raised by Keller and Marian Breland. As you recall, they asked whether all responses are about equally conditionable to all stimuli. In the following article Dobrzecka and his co-workers show which stimuli an animal will use for a given task depends very much on the nature of the task. In a two-response differentiation problem in which a left paw response is required for one stimulus and a right paw response for the other, dogs learn primarily about the direction from which the sound comes and very little about its quality. On the other hand, using the identical stimuli, if the task at hand is a *go-no-go* problem in which the dogs should respond to one stimulus and not respond to a second, then they learn primarily about the quality of the stimulus and only slightly about the direction. Clearly, the assumption that all stimuli are equally available for all responses is wrong.

The abbreviations *Ma* and *Bp* stand for *Metronome anterior* and *Buzzer posterior*, respectively.

QUALITATIVE VERSUS DIRECTIONAL CUES IN TWO FORMS OF DIFFERENTIATION

Czesława Dobrzecka
Genowefa Szwejkowska
Jerzy Konorski

Abstract. Dogs given opportunities to base their instrumental conditioned responses in differentiation learning on either the quality of the auditory conditioned stimulus (for example, metronome versus buzzer) or the direction of its source (in front or behind) choose different cues in different tasks. In $S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$ (left leg-right leg) differentiation they exclusively use directional cues and are almost unable to learn this task when only quality cues are available. When confronted with Pavlovian $S + \rightarrow R$, $S - \rightarrow \text{no } R$ (go-no go) differentiation, however, they generally learn on the basis of quality cues, although some animals also attend to the directional cues. Thus an animal's success or failure in a given differentiation procedure depends not only on its ability to discriminate the stimuli but also on the task with which it is confronted.

Lawicka (1) has shown that, in a free-moving situation, success or failure of training in go left-go right differentiation or go-no go differentiation depends on the character of auditory cues used for the particular task. While for go left-go right differentiation the adequate cues are provided by auditory stimuli presented from different directions, for go-no go differentiation they are provided by stimuli of different quality. We have

now further investigated the same problem, using a different technique.

We used 29 dogs in a Pavlovian soundproof conditioned reflex (CR) chamber. An animal, placed on a stand, was given food by remote control from a feeder situated before him. An instrumental CR consisted in placing the left or right foreleg on the feeder in response to a conditioned stimulus (CS); intertrial intervals were about 1 minute.

In experiment 1, 11 dogs were trained to place their right forelegs on the feeder in response to the sound of a metronome situated in front of

them (Ma) and to place there their left forelegs in response to a buzzer situated behind them (Bp). During training the two stimuli were presented randomly; 5 seconds after presentation of the stimulus the appropriate leg was passively placed on the feeder and food was immediately delivered. Passive placement was remotely controlled by a system of ropes and pulleys. After a few sessions the animals started to execute the trained movements actively in response to the CS's. If the response was correct, food was immediately presented; if incorrect, the CS was discontinued and food was not delivered. If a dog did not respond to 5 seconds of CS, it was prompted by a tug on the proper leg. Eight reinforced trials were given per session, each CS being presented four times in random order.

The task was mastered to a criterion of 80 consecutive correct responses in an average of 230 (range, 120 to 360) trials; these scores include

the trials with passive movements. During each of the ten test sessions that followed, interspersed among the regular trials were two trials with (i) a buzzer presented in front of the animal (Ba) and (ii) a metronome presented behind it (Mp). In all trials either movement was reinforced by food. The results (Table 1) show that in eight dogs at least 90 percent of the responses were determined by the direction of the CS, with total neglect of its quality; that is, the metronome presented from behind evoked the same movement as the buzzer from behind, while the buzzer presented in front evoked the same movement as the metronome in front. In three other dogs the responses were mixed. In no dog, however, did the quality cue prevail over the directional cue.

Further evidence indicating the significance of the directional cues for right leg-left leg differentiation is provided by the results of a control procedure. Six dogs were similarly trained

TABLE 1.

Results of test trials in $S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$ differentiation with 11 dogs, in which the metronome was shifted behind (Mp) and the buzzer in front of (Ba) the animals.

RESPONSES (NO.)					
Mp			Ba		
TO QUALITY (M)	TO DIRECTION (p)	NONE	TO QUALITY (B)	TO DIRECTION (a)	NONE
0	9	1	0	10	0
0	10	0	0	10	0
1	8	1	0	10	0
0	10	0	1	9	0
0	10	0	0	10	0
0	10	0	2	8	0
0	10	0	1	9	0
5	5	0	2	8	0
2	8	0	0	10	0
0	10	0	0	10	0
4	1	5			
<i>Totals (%)</i>					
10.9	82.7	6.4	5.5	94.5	0

except that both CS's sounded from the same point in front of the animal, so that the directional cue was absent; three of them eventually mastered the task after 1000 to 1500 trials, while the other three could not learn even after longer training. In all six dogs symptoms of neurosis developed from time to time. In contrast, when two CS's of different modalities (visual versus auditory) were presented in front of two dogs, the task of right leg-left leg differentiation was quickly learned within about 250 trials.

In experiment 2, ten dogs were trained in go-no go differentiation, the instrumental response being movement of the right foreleg. Response by this movement to the positive CS was always reinforced by food, while the negative CS never brought food. Either the buzzer or a 900-cy/sec tone was the positive CS, while the metronome or a 600-cy/sec tone was the negative CS. For five dogs the positive CS was in front of the animal and the negative one was behind; for the other five these positions were reversed. The

dogs learned the buzzer-metronome differentiation almost immediately; the high tone-low tone differentiation required a few hundred trials.

When the animals were responding correctly in 100 percent of the trials, 20 test trials were given in the same way as in experiment 1—but always without reinforcement, so as not to teach the animal to respond with the movement to the new stimulus combination. The results (Table 2) show that for every dog but one the negative CS placed in the position of the positive CS completely preserved its negative significance—that is, the animals never performed the taught movement in response to it. As for the positive CS placed in the position of the negative CS, it maintained its positive significance in the majority of trials with seven dogs, while negative responses prevailed in three dogs. In all, the animals reacted according to the quality of the CS in 80 percent of the trials and to the direction of the CS in only 20 percent.

So, as Lawicka discovered, animals

TABLE 2.

Results of test trials in $S_1 \rightarrow R$, $S_2 \rightarrow$ no R differentiation with ten dogs. Symbols B, buzzer; M, metronome; T_1 , 900-cy/sec tone; T_2 , 600-cy/sec tone; a, presented in front; p, presented behind.

CS+	REVERSE LOCATION	RESPONSES (NO.) TO		CS—	REVERSE LOCATION	RESPONSES (NO.) TO	
		QUALITY	DIRECTION			QUALITY	DIRECTION
Ba	Bp	0	10	Mp	Ma	10	0
Ba	Bp	3	7	Mp	Ma	9	1
Ba	Bp	9	1	Mp	Ma	10	0
Bp	Ba	4	6	Ma	Mp	10	0
Bp	Ba	10	0	Ma	Mp	10	0
Bp	Ba	7	3	Ma	Mp	10	0
T_1a	T_1p	9	1	T_2p	T_2a	10	0
T_1a	T_1p	8	2	T_2p	T_2a	0	10
T_1p	T_1a	10	0	T_2a	T_2p	10	0
T_1p	T_1a	10	0	T_2a	T_2p	10	0
Totals (%)							
		70	30			89	11

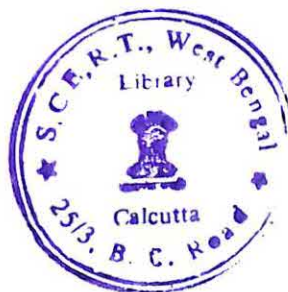
trained in a differentiation procedure requiring two different instrumental responses to two auditory stimuli mainly use directional cues; they are almost unable to learn the task when confronted with purely qualitative cues. On the other hand, in a go-no go differentiation procedure based on reinforcement-versus-nonreinforcement of responses to two auditory stimuli mainly utilize qualitative cues.

These facts have been tentatively interpreted in detail (2). It is notable that monkeys (3) also can establish without difficulty a go-no go differentiation between two different tones

emanating from the same point, while their go left-go right differentiations between these stimuli are as difficult as they are for dogs; in contrast, the go right-go left differentiation between directional cues is easy.

References

1. LAWICKA, W., *Bull. acad. polon. sci. Classe II*, 12, 35 (1964).
2. KONORSKI, J., in *Excerpta Med. Intern. Congr. Ser. 49 Leiden* (1962), pp. 318-29; *Acta Biol. Exp.* 24, 59 (1964).
3. LAWICKA, W., unpublished.



INTRODUCTION

In this reading the authors question the use of a simple stimulus-response connection, or reflex, as a basic unit in the description of behavior. They note, rather, that many activities of both infrahuman animals and men are of the general form "do x until you have done enough." For example, a rat building a nest may shred newspaper until enough newspaper has been shredded to form a nest. A mechanic inflating a tire will add air until his tire gauge reads the appropriate pressure. In cutting a steak one moves a knife back and forth in sequential operations until one observes that a cut has been made completely through the steak. The authors would break down each of these operations into "Test," "Operate," and "Exit" units which are combined in the general structure—Test, Operate, Test, Operate, Test, Operate, . . . Test, Exit. In the steak cutting example the *Test* would amount to asking the question "Has the steak been cut into two pieces?" The *Operation* is one single slicing motion with the knife. An overall description, in this vocabulary, is as follows: Test: Is the steak cut into two pieces? Answer: *No*, therefore Operate: Move the knife back and forth once. Test again: Is the steak cut into two pieces? Answer: *No*, therefore Operate: move the knife back and forth once, and then test again. Continue the (Test, Operate) cycle until the outcome of the test is positive: The steak is now cut into two pieces. At this point, Exit from this "loop" and move on to other matters such as getting the piece that you have cut off into your mouth. The simplest possible cycle of this sort is a *Test, Operate, Test, Exit* cycle, TOTE, from which the authors obtained the name of their basic unit, the TOTE unit. As you read this article think of examples from your own behavior. Do they seem consistent with this basic notion?

THE UNIT OF ANALYSIS

G. A. Miller
E. Galanter
K. Pribram

Most psychologists take it for granted that a scientific account of the behavior of organisms must begin with the definition of fixed, recognizable, elementary units of behavior—something a psychologist can use as a biologist uses cells, or an astronomer uses stars, or a physicist uses atoms, and so on. Given a simple unit, complicated phenomena are then describable as lawful compounds. That is the essence of the highly successful strategy called “scientific analysis.”

The elementary unit that modern, experimental psychologists generally select for their analysis of behavior is the *reflex*. “The isolation of a reflex,” B. F. Skinner tells us, “is the demonstration of a predictable uniformity in behavior. In some form or other it is an inevitable part of any science of behavior. . . . A reflex is not, of course, a theory. It is a fact. It is an analytical unit, which makes the investigation of behavior possible.”¹ Skinner is quite careful to define a re-

flex as a unit of behavior that will yield orderly data: “The appearance of smooth curves in dynamic processes marks a unique point in the progressive restriction of a preparation, and it is to this uniquely determined entity that the term reflex may be assigned.”² This somewhat odd approach to the reflex—in terms of the smoothness of curves—results from Skinner’s consistent attempt to define a unit of behavior in terms of behavior itself instead of by reference to concepts drawn from some other branch of science.

Although Skinner’s approach abolishes the psychologist of certain burdensome responsibilities toward his biological colleagues, the fact remains that the reflex is a concept borrowed originally from physiology and made to seem psychologically substantial largely by the myth of the *reflex arc*: stimulus → receptor → afferent nerve → connective fibers → efferent nerve → effector → response. For many years all those elementary textbooks of psychology that mentioned the nervous system featured the traditional, simplified diagram of the reflex arc in a very prominent position. You may ignore a behaviorist when he tells you that the reflex is a fact,

Reprinted from Plans and the Structure of Behavior by George A. Miller, Eugene Galanter, and Karl H. Pribram. Copyright © 1960 by Holt, Rinehart and Winston, Inc. Reprinted by permission of Holt, Rinehart and Winston, Inc.

¹ B. F. Skinner, *The Behavior of Organisms* (New York: Appleton-Century-Crofts, 1938), p. 9.

² *Ibid.*, p. 40.

but you can scarcely ignore a physiologist when he draws you a picture of it. You might as well deny the small intestines or sneer at the medulla oblongata as to doubt the reflex arc. Even the most obstinate opponent of physiological explanations in psychology can scarcely forget the bloody tissue from which the reflex—even the reflex-sans-arc—originally grew.

But let us suppose, by a wild and irresponsible flight of fancy, that the physiologists and neurologists suddenly announced that they had been mistaken, that there was no such fact as a reflex arc and that the data on which the theory had been based were actually quite different from what had originally been supposed. What then would psychologists say? Would they persist in talking about reflexes? Has the reflex concept been so tremendously helpful that behaviorists could not afford to give it up, even if its biological basis were demolished?

There is some reason to think that the reflex unit has been vastly overrated and that a good many psychologists would like to get out from under it if they could. The reflex arc may have been helpful in getting psychology started along scientific paths, but the suspicion has been growing in recent years that the reflex idea is too simple, the element too elementary. For the most part, serious students of behavior have had to ignore the problem of units entirely. Or they have had to modify their units so drastically for each new set of data that to speak of them as elementary would be the most unblushing sophistry. After watching psychologists struggle under their burden of conditioning reflexes, Chomsky, the linguist and logician, recently summarized their plight in the following terms:

The notions of "stimulus," "response," "reinforcement" are relatively well defined with respect to the barpressing experiments and others similarly restricted. Before we can extend them to real-life behavior, however, certain difficulties must be faced. We must decide, first of all, whether any physical event to which the organism is capable of reacting is to be called a stimulus on a given occasion, or only one to which the organism in fact reacts; and correspondingly, we must decide whether any part of behavior is to be called a response, or only one connected with stimuli in lawful ways. Questions of this sort pose something of a dilemma for the experimental psychologist. If he accepts the broad definitions, characterizing any physical event impinging on the organism as a stimulus and any part of the organism's behavior as a response, he must conclude that behavior has not been demonstrated to be lawful. In the present state of our knowledge, we must attribute an overwhelming influence on actual behavior to ill-defined factors of attention, set, volition, and caprice. If we accept the narrower definitions, then behavior is lawful by definition (if it consists of responses); but this fact is of limited significance, since most of what the animal does will simply not be considered behavior.³

Faced with the choice of being either vague or irrelevant, many psychologists have been restive and ill at ease with their borrowed terms. What went wrong? How was the reflex arc conceived originally, and for what purpose? Can we supplant the reflex arc with some theory of the reflex that is more suited to our current knowledge and interests?

Sir Charles Sherrington and Ivan

³ The passage quoted is from page 30 of Chomsky's review of B. F. Skinner, *Verbal Behavior*, in *Language*, 1959, 35, 26-58.

Petrovitch Pavlov are the two men who are probably most responsible for confirming the psychologist's Image of man as a bundle of S-R reflexes. Yet one may be permitted to speculate that neither of them would approve of the way their concepts have been extended by psychologists. In his *Integrative Action of the Nervous System* (1906) Sherrington is particularly explicit in his qualifications and warnings about the reflex. Again and again he states that "the simple reflex is a useful fiction"—useful for the study of the spinal preparation. He expressed considerable doubt that a stretch reflex, of which the knee jerk is the most frequently quoted example, represented his notion of a simple reflex and questioned whether it should be considered a reflex at all. The synapse was invented by Sherrington in order to explain the differences between the observed properties of nerve trunks and the properties that had to be inferred to describe the neural tissue that intervenes between receptor stimulation and effector response. Nerve trunks will transmit signals in either direction. Characteristically, the signals are of an all-or-none type. Reflex action, on the other hand, is unidirectional and the response is characteristically graded according to the intensity of the stimulus. How can these be reconciled? Sherrington resolved the differences by supporting the neuron doctrine: the nervous system is made up of discrete neural units that have the properties of nerve trunks; intercalated between these units are discontinuities which he christened "synapses," and these have the properties unique to reflexes.

In recent years, graded responses have been shown to be a prepotent characteristic not only of synapses but

also of all excitable tissue, for example, of the finer arborizations of the nerve cells. The cerebral cortex, man's claim to phylogenetic eminence, "still operates largely by means of connections characteristic of primitive neuropil [which is] the most appropriate mechanism for the maintenance of a continuous or steady state, as contrasted to the transmission of information about such states."⁴

Moreover, additional data have come to light. Today we know that neural and receptor tissues are spontaneously active irrespective of environmental excitation. This spontaneous activity is, of course, altered by environmental events—but the change in spontaneous activity may outlast the direct excitation by hours and even days. Furthermore, we know now that the activity of receptors is controlled by efferents leading to them from the central nervous system. As an example, consider the events that control muscular contraction. (Similar, though not identical, mechanisms have also been described for the various sensory systems.) One third of the "motor" nerve fibers that go to muscle actually end in spindles that are the stretch-sensitive receptors. Electrical stimulation of these nerve fibers does not result in contraction of muscle; but the number of signals per unit time that are recorded from the "sensory" nerves coming from the spindles is altered drastically. It is assumed, therefore, that the central nervous mechanism must compare the incoming pattern of signals with the centrally originating "spindle control" signal pattern in order to determine what contribution the muscular con-

⁴ George Bishop, The natural history of the nerve impulse, *Physiological Reviews*, 1956, 36, 376-399.

traction has made to the "spindle sensing" pattern. The outcome of this comparison, or *test*, constitutes the stimulus (the psychophysicist's *proximal* stimulus) to which the organism is sensitive. The test represents the conditions which have to be met before the response will occur. The test may occur in the receptor itself (e.g., in the retina) or in a more centrally located neuronal aggregate (as is probably the case for muscle stretch).

It is clear from examples such as this that the neural mechanism involved in reflex action cannot be diagrammed as a simple reflex arc or even as a chain of stimulus-response connections. A much more complex kind of monitoring, or testing, is involved in reflex action than the classical reflex arc makes any provision for. The only conditions imposed upon the stimulus by the classical chain of elements are the criteria implicit in the thresholds of each element; if the distal stimulus is strong enough to surmount the thresholds all along the arc, then the response must occur. In a sense, the threshold is a kind of test, too, a condition that must be met, but it is a test of strength only. And it must have encouraged psychologists to believe that the only meaningful measurement of a reflex was its strength (probability, magnitude, or latency).

The threshold, however, is only one of many different ways that the input can be tested. Moreover, the response of the effector depends upon the outcome of the test and is most conveniently conceived as an effort to modify the outcome of the test. The action is initiated by an "incongruity" between the state of the organism and the state that is being tested for, and the action persists until the incongruity (i.e., the proximal stimulus) is

removed. The general pattern of reflex action, therefore, is to test the input energies against some criteria established in the organism, to respond if the result of the test is to show an incongruity, and to continue to respond until the incongruity vanishes, at which time the reflex is terminated. Thus, there is "feedback" from the result of the action to the testing phase, and we are confronted by a recursive loop. The simplest kind of diagram to represent this conception of reflex action—an alternative to the classical reflex arc—would have to look something like Figure 1.

The interpretation toward which the argument moves is one that has been called the "cybernetic hypothesis," namely, that the fundamental building block of the nervous system is the feedback loop.⁵ The development of a mathematical theory for servomechanisms, wedded to the physiological accounts of homeostatic mechanisms, has stimulated widespread discussion and speculation about devices closely akin to Figure 1. The argument, therefore, moves toward popular ground.

But what good is this alternative interpretation of the reflex? The psychologist was interested in reflexes because he thought they might provide the units he needed to describe behavior. But simple reflexes have been inadequate. And if reflexes based on afferent-efferent arcs would not turn the trick, why should we hope for better things from reflexes based on

⁵ Norbert Wiener, *Cybernetics* (New York: Wiley, 1948). For a short review of the early development of this idea, see J. O. Wisdom, "The hypothesis of cybernetics," *British Journal for the Philosophy of Science*, 1951, 2, 1-24. For more comprehensive discussion, see W. Sluckin, *Minds and Machines* (London: Penguin, 1954). A little of the cybernetic story is reviewed in the next chapter.

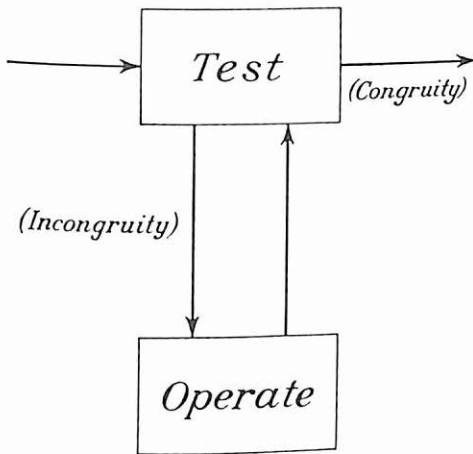


Figure 1

The TOTE unit.

feedback loops? It is the reflex itself—not merely the reflex arc—that has not merely failed, and repairing the neurological theory underlying it is not likely to save the day. What do we hope to gain from such a reinterpretation?

Obviously, the reflex is not the unit we should use as the element of behavior: the unit should be the feedback loop itself. If we think of the Test-Operate-Test-Exit unit—for convenience, we shall call it a TOTE unit—as we do of the reflex arc, in purely anatomical terms, it may describe reflexes, but little else. That is to say, the reflex should be recognized as only one of many possible actualizations of a TOTE pattern. The next task is to generalize the TOTE unit so that it will be useful in a majority—hopefully, in all—of the behavioral descriptions we shall need to make.

Consider what the arrows in Figure 1 might represent. What could flow along them from one box to another? We shall discuss three alternatives: energy, information, and control. If we think of *energy*—neural impulses,

for example—flowing from one place to another over the arrows, then the arrows must correspond to recognizable physical structures—neurons, in the example chosen. As a diagram of energy flow over discrete pathways, therefore, the TOTE unit described in Figure 1 might represent a simple reflex. Or it might represent a servo-mechanism.

There is, however, a second level of abstraction that psychologists usually prefer. We can think of *information* as flowing from one place to another over the arrows. According to the method of measuring information that has been developed by Norbert Wiener and by Claude Shannon, information is transmitted over a channel to the extent that the output of the channel is correlated with the input.⁶ We could therefore think of this second level of abstraction as the transmission of correlation over the arrows. In that case, we are concerned not with the particular structures or kinds of energy that are involved in producing the correlation but only with the fact that events at the two ends of the arrow are correlated. The situation is quite familiar to psychologists, for it is exactly what they mean when they draw an arrow leading from Stimulus to Response in their S-R diagrams or when they define a reflex as a correlation between *S* and *R* but refuse to talk about the neurological basis for that correlation.

A third level of abstraction, how-

⁶ A short introduction to these ideas written for psychologists can be found in G. A. Miller, What is information measurement? *American Psychologist*, 1953, 8, 3-11. A fuller account has been given by Fred Attneave, *Applications of Information Theory to Psychology* (New York: Holt, 1959). See also the highly readable account by Colin Cherry, *On Human Communication* (Cambridge: Technology Press, 1957).

ever, is extremely important for the ideas we shall discuss in the pages that follow. It is the notion that what flows over the arrows in Figure 1 is an intangible something called *control*. Or perhaps we should say that the arrow indicates only succession. This concept appears most frequently in the discussion of computing machines, where the control of the machine's operations passes from one instruction to another, successively, as the machine proceeds to execute the list of instructions that comprise the program it has been given. But the idea is certainly not limited to computers. As a simple example drawn from more familiar activities, imagine that you wanted to look up a particular topic in a certain book in order to see what the author had to say about it. You would open the book to the index and find the topic. Following the entry is a string of numbers. As you look up each page reference in turn, your behavior can be described as under the control of that list of numbers, and control is transferred from one number to the next as you proceed through the list. The transfer of control could be symbolized by drawing arrows from one page number to the next, but the arrows would have a meaning quite different from the two meanings mentioned previously. Here we are not concerned with a flow of energy or transmission of information from one page number to the next but merely with the order in which the "instructions" are executed.

At this abstract level of description we are no longer required to think of the test as a simple threshold that some stimulus energy must exceed. The test phase can be regarded as any process for determining that the operational phase is appropriate. For

example, to be clear though crude, we do not try to take the square root of "ratiocinate." We may know full well how to extract square roots, but before we can execute that operation we must have digits to work on. The operation of extracting square roots is simply irrelevant when we are dealing with words. In order to ensure that an operation is relevant, a test must be built into it. Unless the test gives the appropriate outcome, control cannot be transferred to the operational phase.

When Figure 1 is used in the discussion of a simple reflex it represents all three levels of description simultaneously. When it is used to describe more complex activities, however, we may want to consider only the transfer of information and control or in many instances only the transfer of control. In all cases, however, the existence of a TOTE should indicate that an organizing, coordinating unit has been established, that a Plan is available.

In the following pages we shall use the TOTE as a general description of the control processes involved; the implications it may have for functional anatomy will remain more or less dormant until Chapter 14, at which point we shall indulge in some neuropsychological speculations. Until then, however, the TOTE will serve as a description at only the third, least concrete, level. In its weakest form, the TOTE asserts simply that the operations an organism performs are constantly guided by the outcomes of various tests.

The present authors feel that the TOTE unit, which incorporates the important notion of feedback, is an explanation of behavior in general, and of reflex action in particular, fundamentally different from the explanation provided by the reflex arc.

Consequently, the traditional concepts of stimulus and response must be re-defined and reinterpreted to suit their new context. Stimulus and response must be seen as phases of the organized, coordinated act. We might summarize it this way:

The stimulus is that phase of the forming coordination which represents the conditions which have to be met in bringing it to a successful issue; the response is that phase of one and the same forming coordination which gives the key to meeting these conditions, which serves as instrument in effecting the successful coordination. They are therefore strictly correlative and contemporaneous.⁷

Because stimulus and response are correlative and contemporaneous, the stimulus processes must be thought of not as preceding the response but rather as guiding it to a successful elimination of the incongruity. That is to say, stimulus and response must be considered as aspects of a feedback loop.

The need for some kind of feedback channel in the description of behavior is well recognized by most reflex theorists, but they have introduced it in a peculiar way. For example, it is customary for them to speak of certain consequences of a reflex action as strengthening, or reinforcing, the reflex—such reinforcing consequences of action are a clear example of feedback. Reinforcements are, however, a special kind of feedback that should not be identified with the feedback involved in a TOTE unit. That is to say: (1) a reinforcing feedback must strengthen something, whereas feed-

back in a TOTE is for the purpose of comparison and testing; (2) a reinforcing feedback is considered to be a stimulus (e.g., pellet of food), whereas feedback in a TOTE may be a stimulus, or information (e.g., knowledge of results), or control (e.g., instructions); and (3) a reinforcing feedback is frequently considered to be valuable, or "drive reducing," to the organism, whereas feedback in a TOTE has no such value. . . .

When a TOTE has been executed—the operations performed, the test satisfied, and the exit made—the creature may indeed appear to have attained a more desirable state. It may even be true, on the average, that the TOTE units that are completed successfully in a given situation tend to recur with increased probability, although such a relation would not be necessary. Thus it is possible to discuss a TOTE in the language of reinforcements. Nevertheless, the TOTE involves a much more general conception of feedback. The concept of reinforcement represents an important step forward from reflex arcs toward feedback loops, but bolder strides are needed if behavior theory is to advance beyond the description of simple conditioning experiments.

Perhaps variations in the basic TOTE pattern will prove necessary, so for the purposes of the present discussion we shall continue to regard the diagram in Figure 1 as a hypothesis rather than a fact. The importance of this hypothesis to the general thesis of the book, however, should not be overlooked. It is, in capsule, the account we wish to give of the relation between Image and action. The TOTE represents the basic pattern in which our Plans are cast, the test phase of the TOTE involves the specification of whatever knowledge

⁷ This passage is from an article by John Dewey entitled, "The Reflex Arc Concept in Psychology," an article as valuable today for its wisdom and insight as it was in 1896.

is necessary for the comparison that is to be made, and the operational phase represents what the organism does about it—and what the organism does may often involve overt, observable actions. Figure 1, therefore, rephrases the problem posed in Chapter 1: How does a Plan relate the organism's Image of itself and its universe to the actions, the responses, the behavior that the organism is seen to generate?

Let us see what we must do in order to expand this proposal into something useful. One of the first difficulties—a small one—is to say more exactly what we mean by the “incongruity” that the test phase is looking for. Why not talk simply about the difference, rather than the incongruity, as providing the proximal stimulus? The answer is not profound: We do not want to bother to distinguish between TOTEs in which the operations are performed only when a difference is detected (and where the operations serve to diminish the difference) and TOTEs in which the operations are released only when no difference is detected. When the diagram is used to describe servomechanisms, for example, it is quite important to distinguish “positive” from “negative” feedback, but, because we are going to be interested primarily in the feedback of control, such questions are not critical. Rather than treat all these varieties as different units of analysis, it seems simpler to treat them all as examples of a more general type of “incongruity-sensitive” mechanism.⁸

⁸ The notion of an “incongruity-sensitive” mechanism appears to the authors to be related to Festinger's conceptions of “cognitive dissonance,” but we have not attempted to explore or develop that possibility. See Leon Festinger, *A Theory of Cognitive Dissonance* (Evanston: Row, Peterson, 1957).

A second difficulty—this one rather more important—is the question of how we can integrate this TOTE unit into the sort of hierarchical structure of behavior that we insisted on in Chapter 1. How can the two concepts—feedback and hierarchy—be reconciled? One method of combining feedback components in a hierarchy has been described by D. M. MacKay,⁹ who proposed to make the consequences of the operational phase in one component provide the input to the comparator of a second component; MacKay's suggestion leads to a string of such feedback components, each representing a progressively higher degree of abstraction from the external reality. Although MacKay's scheme is quite ingenious, we are persuaded that a somewhat different method of constructing the hierarchy will better serve a psychologist's descriptive purposes. A central notion of the method followed in these pages is that the operational components of TOTE units may themselves be TOTE units. That is to say, the TOTE pattern describes both strategic and tactical units of behavior. Thus the operational phase of a higher-order TOTE might itself consist of a string of other TOTE units, and each of these, in turn, may contain still other strings of TOTEs, and so on. Since this method of retaining the same pattern of description for the higher, more strategic units as for the lower, more tactical units may be confusing on first acquaintance, we shall consider an example.

R. S. Woodworth has pointed out how frequently behavioral activities

⁹ D. M. MacKay, The epistemological problem for automata, in C. E. Shannon and J. McCarthy, eds., *Automata Studies* (Princeton: Princeton University Press, 1956), pp. 235–251.

are organized in two stages.¹⁰ Woodworth refers to them as "two-phase motor units." The first phase is preparatory or mobilizing; the second, effective or consummatory. To jump, you first flex the hips and knees, then extend them forcefully; the crouch prepares for the jump. To grasp an object, the first phase is to open your hand, the second is to close it around the object. You must open your mouth before you can bite. You must draw back your arm before you can strike, etc. The two phases are quite different movements, yet they are obviously executed as a single unit of action. If stimulation is correct for releasing the action, first the preparatory TOTE unit is executed, and when it has been completed the stimulation is adequate for the consummatory TOTE unit and the action is executed. Many of these two-phase plans are repetitive: the completion of the second phase in turn provides stimuli indicating that the execution of the first phase is again possible, so an alternation between the two phases is set up, as in walking, running, chewing, drinking, sweeping, knitting, etc.

We should note well the construction of a "two-phase" TOTE unit out of two simpler TOTE units. Consider hammering a nail as an example. As a Plan, of course, hammering has two phases, lifting the hammer and then striking the nail. We could represent it by a tree, or hierarchy, as in Figure 2. If we ask about details, however, the representation of hammering in Figure 2 as a simple list containing two items is certainly too sketchy. It does not tell us, for one thing, how long to go on hammering. What is the "stop rule"? For this, we must

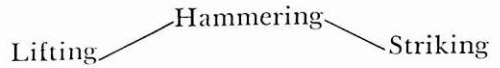


Figure 2

Hammering as a hierarchy.

indicate the test phase, as in Figure 3. The diagram in Figure 3 should indicate that when control is transferred to the TOTE unit that we are calling "hammering," the hammering continues until the head of the nail is flush with the surface of the work. When the test indicates that the nail is driven in, control is transferred elsewhere. Now, however, we seem to have lost the hierarchical structure. The hierarchy is recovered when we look at the box labeled "hammer," for there we find two TOTE units, each with its own test, as indicated in Figure 4. When the pair of TOTE units combined in Figure 4 are put inside the operational phase in Figure 3, the result is the hierarchical Plan for hammering nails that is shown in Figure 5.

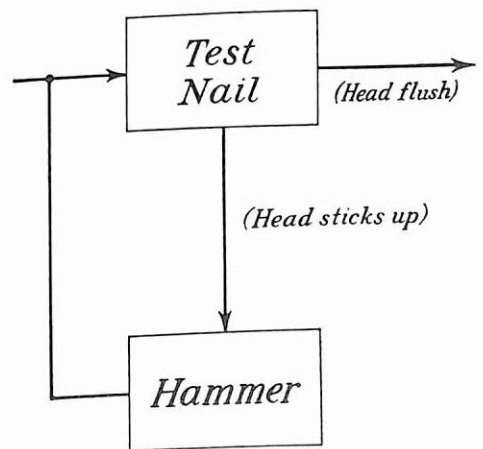


Figure 3

Hammering as a TOTE unit.

¹⁰ Robert S. Woodworth, *Dynamics of Behavior* (New York: Holt, 1958), pp. 36 ff.

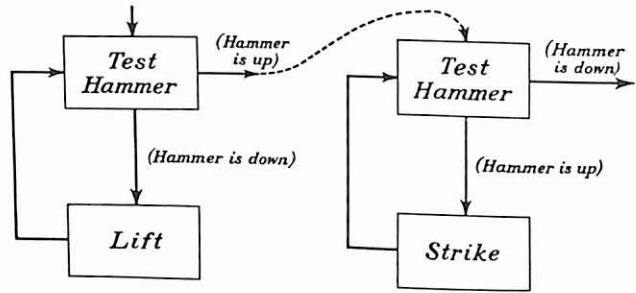


Figure 4

Dashed line indicates how two simple TOTE units are connected to form the operational phase of the more complicated TOTE unit in Figure 3.

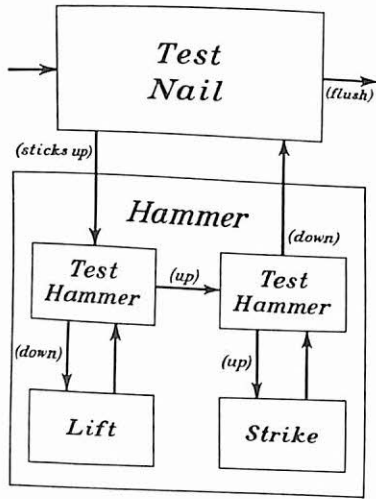


Figure 5

The hierarchical plan for hammering nails.

the test of the nail reveals that its head is flush with the surface of the work, at which point control can be transferred elsewhere. Thus the compound of TOTE units unravels itself simply enough into a coordinated sequence of tests and actions, although the underlying structure that organizes and coordinates the behavior is itself hierarchical, not sequential.

It may seem slightly absurd to analyze the motions involved in hammering a nail in this explicit way, but it is better to amuse a reader than to confuse him. It is merely an illustration of how several simple TOTE units, each with its own test-operate-test loop, can be embedded in the operational phase of a larger unit with its particular test-operate-test loop. Without such an explicit illustration it might not have been immediately obvious how these circles within circles could yield hierarchical trees.

More complicated Plans—Woodworth refers to them as “polyphase motor units”—can be similarly described as TOTE units built up of subplans that are themselves TOTE units. A bird will take off, make a few wing strokes, glide, brake with its wings, thrust its feet forward, and

If this description of hammering is correct, we should expect the sequence of events to run off in this order: Test nail. (Head sticks up.) Test hammer. (Hammer is down.) Lift hammer. Test hammer. (Hammer is up.) Strike nail. Test hammer. (Hammer is down.) Test nail. (Head sticks up.) Test hammer. And so on, until

land on the limb. The whole action is initiated as a unit, is controlled by a single Plan, yet is composed of several phases, each involving its own Plan, which may in turn be comprised of subplans, etc.

Note that it is the *operational* phase of the TOTE that is expanded into a list of other TOTE units. If we wish to preserve the TOTE pattern as it is diagrammed in Figure 1, we cannot use it to build up more complicated tests.¹¹ The tests that are available, therefore, are conceived to be relatively fixed; it is the operational phase that may be quite various and complex. The operational phase may, of course, consist of a list of TOTEs, or it may terminate in efferent activity.¹²

¹¹ The reason that the TOTE of Figure 1 can be expanded only in its operational phase is purely formal and can be appreciated by simple counting: There are four arrows associated with Test; there are two arrows associated with Operate; and there are two arrows associated with TOTE as a unit. Therefore, if the two-arrowed TOTE is used to construct some component of another TOTE, the component it constructs must be the two-arrowed Operate, not the four-arrowed Test. However, rigid restrictions such as these are probably unrealistic and justifiable only in terms of pedagogic simplification. Anyone who has studied the hierarchically organized programs written by Newell, Shaw, and Simon to stimulate human problem-solving will recognize how primitive and unelaborated these TOTE hierarchies really are.

¹² If we take seriously the suggested form of the TOTE, the system may be easily trapped into loops. For example, if the subtests in the expansion of an operational phase all pass, but the basic test fails, a loop will exist. In order to avoid loops we might insist that the basic test imply the disjunction of the subtests. A more realistic solution would accept the occurrence of loops as a signal that the Plan was not successful in producing the result for which the basic test was designed; it would then be necessary to provide further machinery for discovering and stopping such loops.

If we consider complex Plans—TOTE hierarchies in which the operation of one TOTE is itself a list of TOTE units—then some general properties of such systems become apparent:

—The hierarchical structure underlying behavior is taken into account in a way that can be simply described with the computer language developed by Newell, Shaw, and Simon for processing lists.

—Planning can be thought of as constructing a list of tests to perform. When we have a clear Image of a desired outcome, we can use it to provide the conditions for which we must test, and those tests, when arranged in sequence, provide a crude strategy for a possible Plan. (Perhaps it would be more helpful to say that the conditions for which we must test *are* an Image of the desired outcome.)

—The operational phase can contain both tests and operations. Therefore the execution of a Plan of any complexity must involve many more tests than actions. This design feature would account for the general degradation of information that occurs whenever a human being is used as a communication channel.

In lower animals it appears that the pattern of their behavior is normally constructed more or less fortuitously by the environment—only man cherishes the illusion of being master of his fate. That is to say, the environment provides stimuli that “release” the next stage of the animal’s activity. It is something of a philosophical question as to whether we wish to believe in plans that exist somewhere outside of nervous systems, so perhaps we should say merely that lower animals appear to have more tactics than strategy.

As we ascend the evolutionary scale

we find in mammals an increasing complexity in the kind of tests the animals can perform. In man we have a unique capacity for creating and manipulating symbols, and when that versatility is used to assign names to TOTE units, it becomes possible for him to use language in order to rearrange the symbols and to form new

Plans. We have every reason to believe that man's verbal abilities are very intimately related to his planning abilities. And, because human Plans are so often verbal, they can be communicated, a fact of crucial importance in the evolution of our social adjustments to one another. . . .

INTRODUCTION

In this selection, which has considerable charm, Tolman directly tackles the structure of knowledge problem. He looks at the “telephone switchboard” theory that all behavior is made up of responses directly attached to stimuli and more specifically at the particular stimulus-response theory which assumes that the connection between stimulus and response is strengthened or “reinforced” by reward which arrives after the response.

Tolman was an early and effective critic of the very simplest forms of stimulus-response psychology. In the following work he describes a number of experiments which led him to believe that something considerably more complicated—which he called the building of cognitive maps—was going on. As you read try not only to understand Tolman’s criticisms of simple forms of S-R theory, but also try to see the basic problem presented by Tolman’s ideas: What is a cognitive map? How is one formed? When we say that a rat is using a cognitive map, are we saying anything more than that he is behaving about the way that we, as intelligent people, would behave in that situation. If Tolman were saying no more than this then his explanation would not help much for understanding anybody’s behavior.

Try not to be put off by the bits of jargon in the paper. “VTE,” meaning Vicarious Trial and Error, refers simply to the activity of standing at a choice point and looking back and forth at the various alternatives presented. One wonders why such phrases have crept into psychologists’ writing. Perhaps a controlling factor is that “looking back and forth” doesn’t sound as scientific as “Vicarious Trial and Error.”

COGNITIVE MAPS IN RATS AND MEN

Edward C. Tolman

I shall devote the body of this paper to a description of experiments with rats. But I shall also attempt in a few words at the close to indicate the significance of these findings on rats for the clinical behavior of men. Most of the rat investigations, which I shall report, were carried out in the Berkeley laboratory. But I shall also include, occasionally, accounts of the behavior of non-Berkeley rats who obviously have misspent their lives in out-of-State laboratories. Furthermore, in reporting our Berkeley experiments I shall have to omit a very great many. The ones I *shall* talk about were carried out by graduate students (or underpaid research assistants) who, supposedly, got some of their ideas from me. And a few, though a very few, were even carried out by me myself.

Let me begin by presenting diagrams for a couple of typical mazes, an alley maze and an elevated maze. In the typical experiment a hungry rat is put at the entrance of the maze (alley or elevated), and wanders about through the various true path segments and blind alleys until he finally comes to the food box and eats. This

is repeated (again in the typical experiment) one trial every 24 hours and the animal tends to make fewer and fewer errors (that is, blind-alley entrances) and to take less and less time between start and goal-box until finally he is entering no blinds at all and running in a very few seconds from start to goal. The results are usually presented in the form of average curves of blind-entrances, or of seconds from start to finish, for groups of rats.

All students agree as to the facts. They disagree, however, on theory and explanation.

(1) First, there is a school of animal psychologists which believes that the maze behavior of rats is a matter of mere simple stimulus-response connections. Learning, according to them, consists in the strengthening of some of these connections and in the weakening of others. According to this 'stimulus-response' school the rat in progressing down the maze is helplessly responding to a succession of external stimuli—sights, sounds, smells, pressures, etc. impinging upon his external sense organs—plus internal stimuli coming from the viscera and from the skeletal muscles. These external and internal stimuli call out the walkings, runnings, turnings, retracings, smellings, rearings, and the

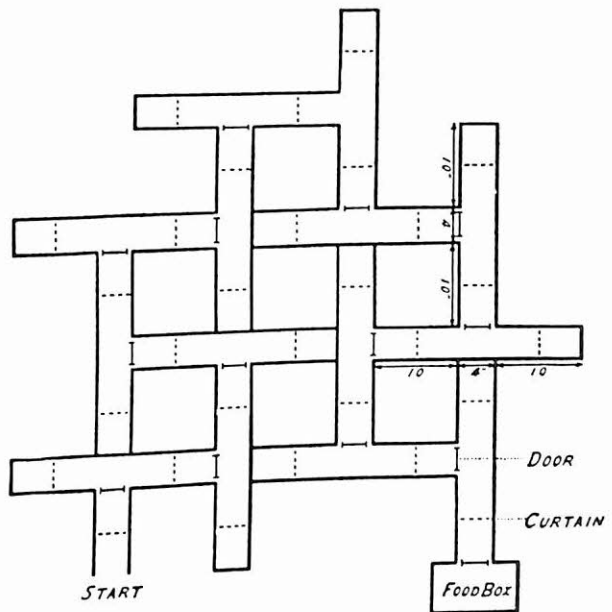


Figure 1

Plan of maze: 14-Unit T-Alley Maze (From M. H. Elliott, The effect of change of reward on the maze performance of rats. *Univ. of Calif. Publ. Psychol.*, 1928, 4, p. 20.)

like which appear. The rat's central nervous system, according to this view, may be likened to a complicated telephone switchboard. There are the incoming calls from sense-organs and there are the outgoing messages to muscles. Before the learning of a specific maze, the connecting switches (synapses according to the physiologist) are closed in one set of ways and produce the primarily exploratory responses which appear in the early trials. *Learning*, according to this view, consists in the respective strengthening and weakening of various of these connections; those connections which result in the animal's going down the true path become relatively more open to the passage of nervous impulses, whereas those which

lead him into the blinds become relatively less open.

It must be noted in addition, however, that this stimulus-response school divides further into two subgroups.

(a) There is a subgroup which holds that the mere mechanics involved in the running of a maze is such that the crucial stimuli from the maze get presented simultaneously with the correct responses more frequently than they do with any of the incorrect responses. Hence, just on a basis of this greater frequency, the neural connections between the crucial stimuli and the correct responses will tend, it is said, to get strengthened at the expense of the incorrect connections.

(b) There is a second subgroup in this stimulus-response school which

holds that the reason the appropriate connections get strengthened relatively to the inappropriate ones is, rather, the fact that the responses resulting from the correct connections are followed more closely in time by need-reductions. Thus a hungry rat in a maze tends to get to food and have his hunger reduced *sooner* as a result of the true path responses than as a result of the blind alley responses. And such immediately following need-reductions or, to use another term, such 'positive reinforcements' tend somehow, it is said, to strengthen the connections which have most closely preceded them. Thus it is as if—although this is certainly not the way this subgroup would themselves state it—the satisfaction-receiving part of the rat telephoned back to Central and said to the girl: "Hold that connection; it was good; and see to it

that you blankety-blank well use it again the next time these same stimuli come in." These theorists also assume (at least some of them do some of the time) that, if bad results—'annoyances,' 'negative reinforcements'—follow, then this same satisfaction-and-annoyance-receiving part of the rat will telephone back and say, "Break that connection and don't you dare use it next time either."

So much for a brief summary of the two subvarieties of the 'stimulus-response,' or telephone switchboard school.

(2) Let us turn now to the second main school. This group (and I belong to them) may be called the field theorists. We believe that in the course of learning something like a field map of the environment gets established in the rat's brain. We agree with the other school that the rat in

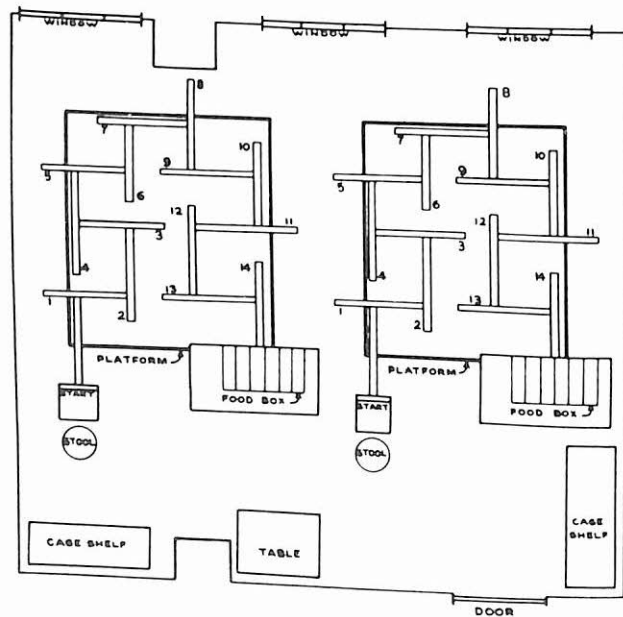


Figure 2
14-Unit T-Elevated Mazes (From C. H. Honzik, *The sensory basis of maze learning in rats. Compar. Psychol. Monogr.*, 1936, 13, No. 4, p. 4. These were two identical mazes placed side by side in the same room.)

running a maze is exposed to stimuli and is finally led as a result of these stimuli to the responses which actually occur. We feel, however, that the intervening brain processes are more complicated, more patterned and often, pragmatically speaking, more autonomous than do the stimulus-response psychologists. Although we admit that the rat is bombarded by stimuli, we hold that his nervous system is surprisingly selective as to which of these stimuli it will let in at any given time.

Secondly, we assert that the central office itself is far more like a map control room than it is like an old-fashioned telephone exchange. The stimuli, which are allowed in, are not connected by just simple one-to-one switches to the outgoing responses. Rather, the incoming impulses are

usually worked over and elaborated in the central control room into a tentative, cognitive-like map of the environment. And it is this tentative map, indicating routes and paths and environmental relationships, which finally determines what responses, if any, the animal will finally release.

Finally, I, personally, would hold further that it is also important to discover in how far these maps are relatively narrow and strip-like or relatively broad and comprehensive. Both strip-maps and comprehensive-maps may be either correct or incorrect in the sense that they may (or may not), when acted upon, lead successfully to the animal's goal. The differences between such strip maps and such comprehensive maps will appear only when the rat is later presented with some change within the given

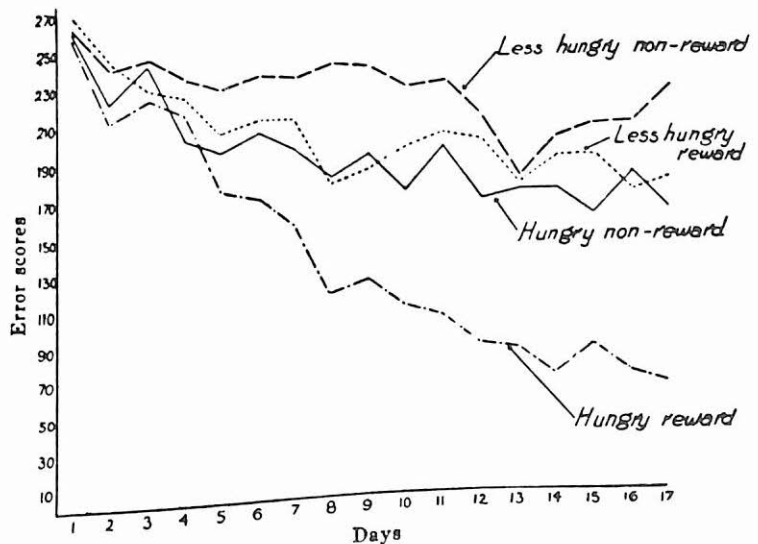


Figure 3

Error curves for four groups, 36 rats (From E. C. Tolman and C. H. Honzik, Degrees of hunger, reward and nonreward, and maze learning in rats. *Univ. Calif. Publ. Psychol.*, 1930, 4, No. 16, p. 246. A maze identical with the alley maze shown in Figure 1 was used.)

environment. Then, the narrower and more strip-like the original map, the less will it carry over successfully to the new problem; whereas, the wider and the more comprehensive it was, the more adequately it will serve in the new set-up. In a strip-map the given position of the animal is connected by only a relatively simple and single path to the position of the goal. In a comprehensive-map a wider arc of the environment is represented, so that, if the starting position of the animal be changed or variations in the specific routes be introduced, this wider map will allow the animal still to behave relatively correctly and to choose the appropriate new route.

But let us turn, now, to the actual experiments. The ones, out of many, which I have selected to report are simply ones which seem especially important in reinforcing the theoretical position I have been presenting. This position, I repeat, contains two assumptions: First, that learning con-

sists not in stimulus-response connections but in the building up in the nervous system of sets which function like cognitive maps, and second, that such cognitive maps may be usefully characterized as varying from a narrow strip variety to a broader comprehensive variety.

The experiments fall under five heads: (1) "latent learning," (2) "vicarious trial and error" or "VTE," (3) "searching for the stimulus," (4) "hypotheses" and (5) "spatial orientation."

(1) "Latent Learning" Experiments

The first of the latent learning experiments was performed at Berkeley by Blodgett. It was published in 1929. Blodgett not only performed the experiments, he also originated the concept. He ran three groups of rats through a six-unit alley maze, shown in Fig. 4. He had a control group and

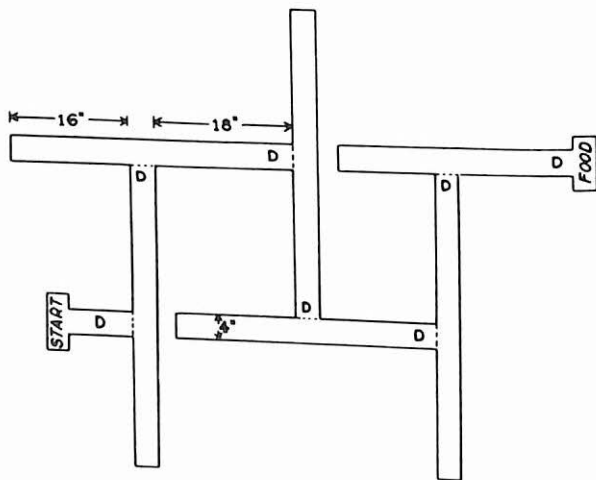


Figure 4
6-Unit Alley T-Maze (From H. C. Blodgett, The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, 4, No. 8, p. 117.)

two experimental groups. The error curves for these groups appear in Fig. 5. The solid line shows the error curve for Group I, the control group. These animals were run in orthodox fashion. That is, they were run one trial a day and found food in the goal-box at the end of each trial. Groups II and III were the experimental groups. The animals of Group II, the dash line, were not fed in the maze for the first six days but only in their home cages some two hours later. On the seventh day (indicated by the small cross) the rats found food at the end of the maze for the first time and continued to find it on subsequent days. The animals of Group III were treated similarly except that they first found food at the end of the maze on the third day and continued to find it there on subsequent days. It will be observed that the experimental

groups as long as they were not finding food did not appear to learn much. (Their error curves did not drop.) But on the days immediately succeeding their first finding of the food their error curves did drop astoundingly. It appeared, in short, that during the non-rewarded trials these animals had been learning much more than they had exhibited. This learning, which did not manifest itself until after the food had been introduced, Blodgett called "latent learning." Interpreting these results anthropomorphically, we would say that as long as the animals were not getting any food at the end of the maze they continued to take their time in going through it—they continued to enter many blinds. Once, however, they knew they were to get food, they demonstrated that during these preceding non-rewarded trials they had learned where many of the blinds were. They had been building

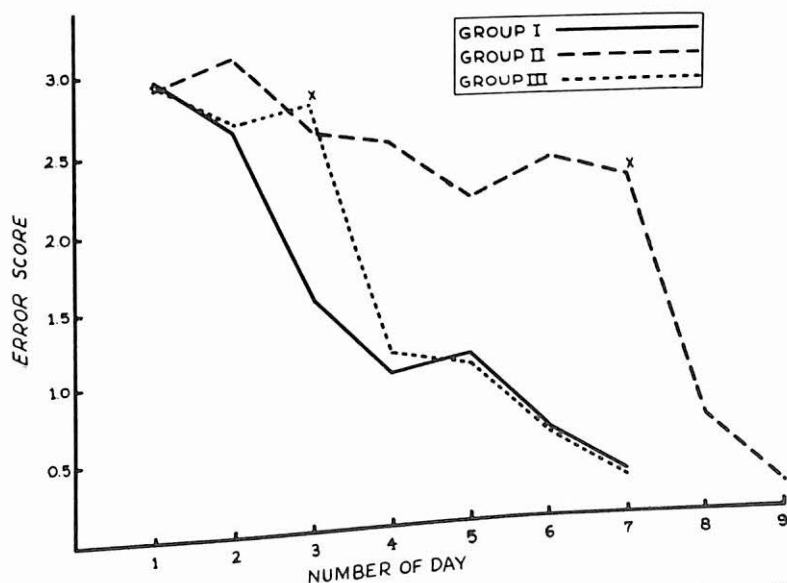


Figure 5

(From H. C. Blodgett, The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, 4, No. 8, p. 120.)

up a 'map,' and could utilize the latter as soon as they were motivated to do so.

Honzik and myself repeated the experiments (or rather he did and I got some of the credit) with the 14-unit T-mazes shown in Fig. 1, and with larger groups of animals, and got similar results. The resulting curves are shown in Fig. 6. We used two control groups—one that never found food in the maze (HNR) and one that found it throughout (HR). The experimental group (HNR-R) found food at the end of the maze from the 11th day on and showed the same sort of a sudden drop.

But probably the best experiment demonstrating latent learning was, unfortunately, done not in Berkeley but at the University of Iowa, by Spence and Lippitt. Only an abstract

of this experiment has as yet been published. However, Spence has sent a preliminary manuscript from which the following account is summarized. A simple Y-maze (see Fig. 7) with two goal-boxes was used. Water was at the end of the right arm of the Y and food at the end of the left arm. During the training period the rats were run neither hungry nor thirsty. They were satiated for both food and water before each day's trials. However, they were willing to run because after each run they were taken out of whichever end box they had got to and put into a living cage, with other animals in it. They were given four trials a day in this fashion for seven days, two trials to the right and two to the left.

In the crucial test the animals were divided into two subgroups one made solely hungry and one solely thirsty.

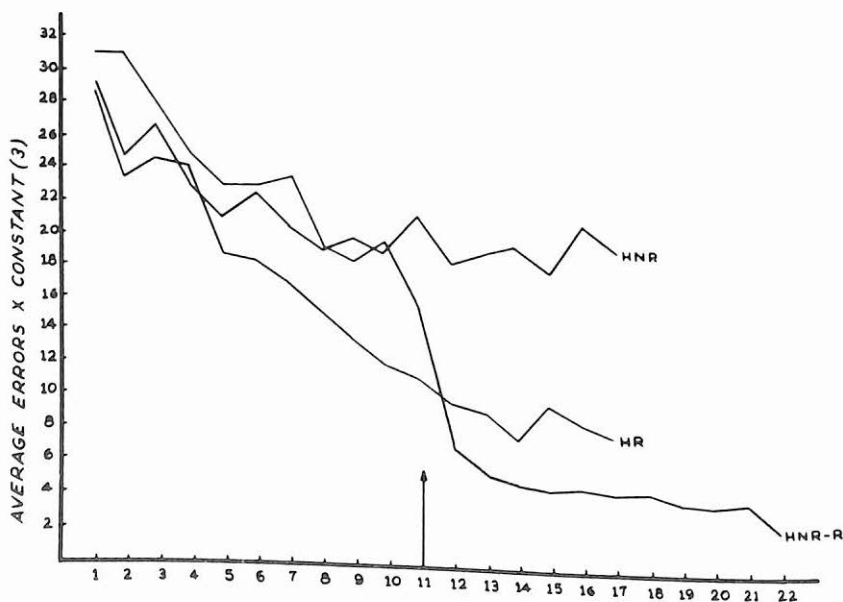


Figure 6

Error curves for HR, HNR, and HNR-R (From E. C. Tolman and C. H. Honzik, Introduction and removal of reward and maze performance in rats, *Univ. Calif. Publ. Psychol.*, 1930, 4, No. 19, p. 267.)

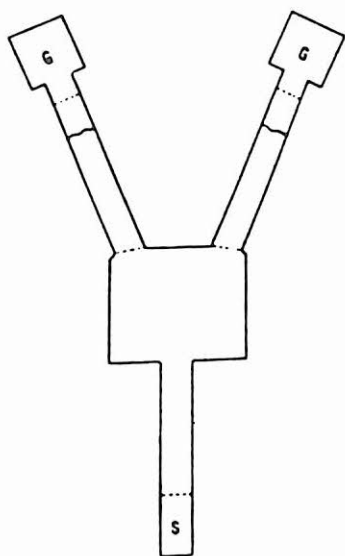


Figure 7

Ground plan of the apparatus (Taken from K. W. Spence and R. Lippitt, An experimental test of the sign-gestalt theory of trial and error learning. *J. exper. Psychol.*, 1946 36, p. 494. In this article they were describing another experiment but used the same maze.)

It was then found that on the first trial the hungry group went at once to the left, where the food had been, statistically more frequently than to the right; and the thirsty group went to the right, where the water had been, statistically more frequently than to the left. These results indicated that under the previous non-differential and very mild rewarding conditions of merely being returned to the home cages the animals had nevertheless been learning where the water was and where the food was. In short, they had acquired a cogni-

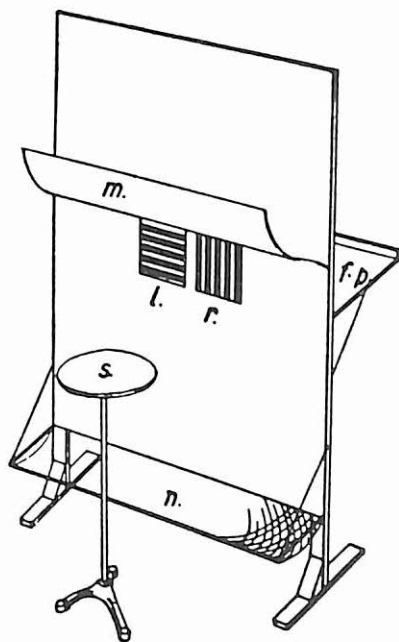


Figure 8

Apparatus Used for Testing Discrimination of Visual Patterns (From K. S. Lashley, The mechanism of vision. I A method for rapid analysis of pattern-vision in the rat. *J. genet. Psychol.*, 1930, 37, p. 454.

tive map to the effect that food was to the left and water to the right, although during the acquisition of this map they had not exhibited any stimulus-response propensities to go more to the side which became later the side of the appropriate goal.

There have been numerous other latent learning experiments done in the Berkeley laboratory and elsewhere. In general, they have for the most part all confirmed the above sort of findings.

Let us turn now to the second group of experiments.

(2) "Vicarious Trial and Error" or "VTE"

The term Vicarious Trial and Error (abbreviated as VTE) was invented by Prof. Muenzinger at Colorado¹ to designate the hesitating, looking-back-and-forth, sort of behavior which rats can often be observed to indulge in at a choice-point before actually going one way or the other.

Quite a number of experiments upon VTEing have been carried out in our laboratory. I shall report only a few. In most of them what is called a discrimination set-up has been used. In one characteristic type of visual discrimination apparatus designed by Lashley (shown in Fig. 8) the animal

¹ Vide: K. F. Muenzinger, Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *J. genet. Psychol.*, 1938, 53, 75-86.

is put on a jumping stand and faced with two doors which differ in some visual property say, as here shown, vertical stripes vs. horizontal stripes.

One of each such pair of visual stimuli is made always correct and the other wrong; and the two are interchanged from side to side in random fashion. The animal is required to learn, say, that the vertically striped door is always the correct one. If he jumps to it, the door falls open and he gets to food on a platform behind. If, on the other hand, he jumps incorrectly, he finds the door locked and falls into a net some two feet below from which he is picked up and started over again.

Using a similar set-up (see Fig. 9), but with landing platforms in front of the doors so that if the rat chose incorrectly he could jump back again and start over, I found that when the

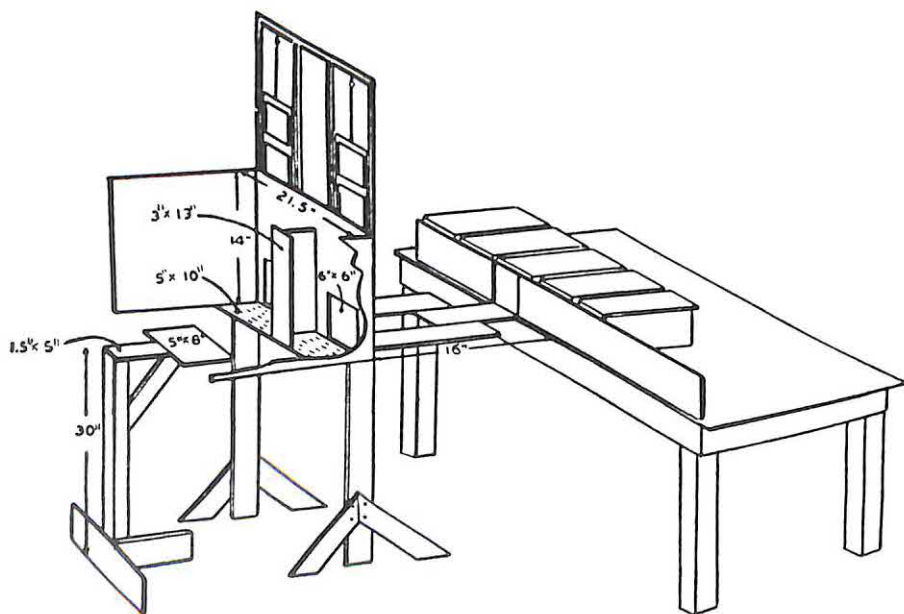


Figure 9
(From E. C. Tolman, Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.*, 1939, 46, p. 319.)

choice was an easy one, say between a white door and a black door, the animals not only learned sooner but also did more VTEing than when the choice was difficult, say between a white door and a gray door (see Fig. 10). It appeared further (see Fig. 11) that the VTEing began to appear just as (or just before) the rats began to learn. After the learning had become established, however, the VTE's began to go down. Further, in a study of individual differences by myself, Geier and Levin² (actually done by Geier and Levin) using this same visual discrimination apparatus, it was found that with one and the same difficulty

of problem the smarter animal did the more VTEing.

To sum up, in *visual discrimination* experiments the better the learning, the more the VTE's. But this seems contrary to what we would perhaps have expected. We ourselves would expect to do more VTEing, more sampling of the two stimuli, when it is difficult to choose between them than when it is easy.

What is the explanation? The answer lies, I believe, in the fact that the manner in which we set the visual discrimination problems for the rats and the manner in which we set similar problems for ourselves are different. We already have our 'instructions.' We know beforehand what it is we are to do. We are told, or we tell ourselves, that it is the lighter of the two grays, the behavior of the

² F. M. Geier, M. Levin & E. C. Tolman, Individual differences in emotionality, hypothesis formation, vicarious trial and error and visual discrimination learning in rats. *Compar. Psychol. Monogr.*, 1941, 17, No. 3.

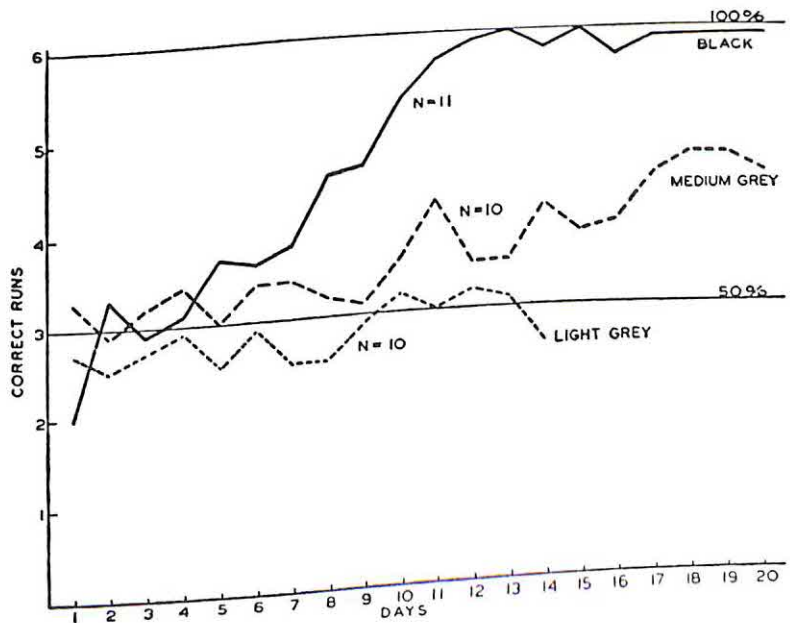


Figure 10

(From E. C. Tolman, Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.*, 1939, 46, p. 319.)

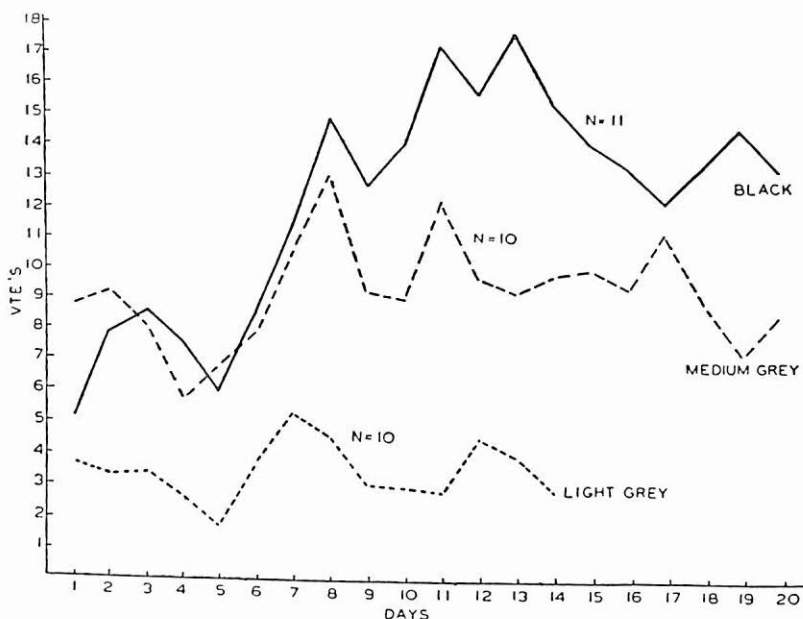


Figure 11

(From E. C. Tolman, Prediction of vicarious trial and error by means of the schematic sowbug. *Psychol. Rev.*, 1939, 46, p. 320.)

two weights, or the like, which is to be chosen. In such a setting we do more sampling, more VTEing, when the stimulus-difference is small. But for the rats the usual problem in a discrimination apparatus is quite different. They do not know what is wanted of them. The major part of their learning in most such experiments seems to consist in their discovering the instructions. The rats have to discover that it is the differences in visual brightness, not the differences between left and right, which they are to pay attention to. Their VTEing appears when they begin to 'catch on.' The greater the difference between the two stimuli, the more the animals are attracted by this difference. Hence the sooner they catch on, and during this catching on, the more they VTE.

That this is a reasonable interpretation appeared further, from an experiment by myself and Minium (the actual work done, of course, by Minium) in which a group of six rats was first taught a white vs. black discrimination, then two successively more difficult gray vs. black discriminations. For each difficulty the rats were given a long series of further trials beyond the points at which they had learned. Comparing the beginning of each of these three difficulties the results were that the rats did more VTEing for the easy discriminations than for the more difficult ones. When, however, it came to a comparison of amounts of VTEing during the final performance after each learning had reached a plateau, the opposite results were obtained. In other words, after the rats had finally

divined their instructions, then they, like human beings, did more VTEing, more sampling, the more difficult the discrimination.

Finally, now let us note that it was also found at Berkeley by Jackson³ that in a maze the difficult maze units produce more VTEing and also that the more stupid rats do the more VTEing. The explanation, as I see it, is that, in the case of mazes, rats know their instructions. For them it is natural to expect that the same spatial path will always lead to the same outcome. Rats in mazes don't have to be told.

But what, now, is the final significance of all this VTEing? How do these facts about VTEing affect our theoretical argument? My answer is that these facts lend further support to the doctrine of a building up of

maps. VTEing, as I see it, is evidence that in the critical stages—whether in the first picking up of the instructions or in the later making sure of which stimulus is which—the animal's activity is not just one of responding passively to discrete stimuli, but rather one of the active selecting and comparing of stimuli. This brings me then to the third type of experiment.

(3) "Searching for the Stimulus"

I refer to a recent, and it seems to me extremely important experiment, done for a Ph.D. dissertation by Hudson. Hudson was first interested in the question of whether or not rats could learn an avoidance reaction in one trial. His animals were tested one at a time in a living cage (see Fig. 13) with a small striped visual pattern at the end, on which was mounted a food cup. The hungry rat approached this food cup and ate. An electrical

³ L. L. Jackson, V. T. E. on an elevated maze. *J. comp. Psychol.*, 1943, 36, 99-107.

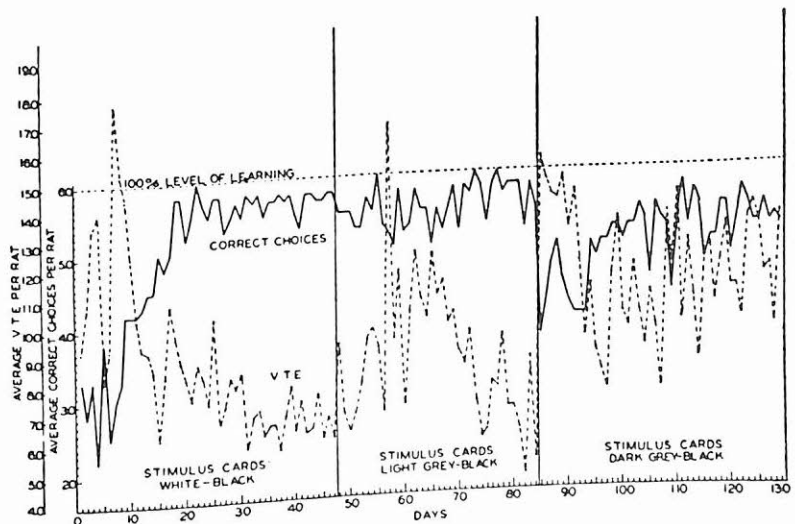


Figure 12

(From E. C. Tolman and E. Minium, VTE in rats: overlearning and difficulty of discrimination. *J. comp. Psychol.*, 1942, 34, p. 303.)

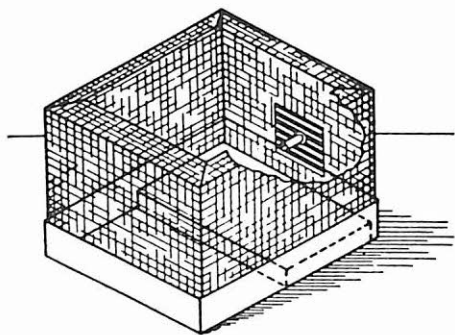


Figure 13

(From Bradford Hudson. Ph.D. Thesis: 'One trial learning: A study of the avoidance behavior of the rat.' On deposit in the Library of the University of California, Berkeley, California.)

arrangement was provided so that when the rat touched the cup he could be given an electric shock. And one such shock did appear to be enough. For when the rat was replaced in this same cage days or even weeks afterwards, he usually demonstrated immediately strong avoidance reactions to the visual pattern. The animal withdrew from that end of the cage, or piled up sawdust and covered the pattern, or showed various other amusing responses all of which were in the nature of withdrawing from the pattern or making it disappear.

But the particular finding which I am interested in now appeared as a result of a modification of this standard procedure. Hudson noticed that the animals, anthropomorphically speaking, often seemed to look around *after* the shock to see what it was that had hit them. Hence it occurred to him that, if the pattern were made to disappear the instant the shock occurred, the rats might not establish the association. And this indeed is what happened in the case of many in-

dividuals. Hudson added further electrical connections so that when the shock was received during the eating, the lights went out, the pattern and the food cup dropped out of sight, and the lights came on again all within the matter of a second. When such animals were again put in the cage 24 hours later, a large percentage showed no avoidance of the pattern. Or to quote Hudson's own words:

Learning what object to avoid . . . may occur exclusively during the period *after* the shock. For if the object from which the shock was actually received is removed at the moment of the shock, a significant number of animals fail to learn to avoid it, some selecting other features in the environment for avoidance, and others avoiding nothing.

In other words, I feel that this experiment reinforces the notion of the largely active selective character in the rat's building up of his cognitive map. He often has to look actively for the significant stimuli in order to form his map and does not merely passively receive and react to all the stimuli which are physically present.

Turn now to the fourth type of experiment.

(4) The "Hypothesis" Experiments

Both the notion of hypotheses in rats and the design of the experiments to demonstrate such hypotheses are to be credited to Krech. Krech used a four-compartment discrimination-box. In such a four-choice box the correct door at each choice-point may be determined by the experimenter in terms of its being lighted or dark, left or right, or various combinations of

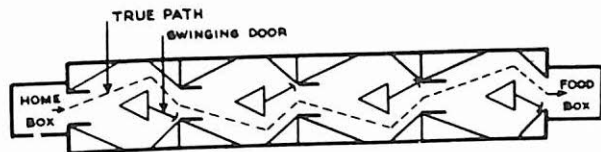


Figure 14

(From I. Krechevsky (Now D. Krech), The genesis of "hypotheses" in rats. *Univ. Calif. Publ. Psychol.*, 1932, 6, No. 4, p. 46.)

these. If all possibilities are randomized for the 40 choices made in 10 runs of each day's test, the problem could be made insoluble.

When this was done, Krech found that the individual rat went through a succession of systematic choices. That is, the individual animal might perhaps begin by choosing practically all right-hand doors, then he might give this up for choosing practically all left-hand doors, and then, for choosing all dark doors, and so on. These relatively persistent, and well-above-chance systematic types of choice Krech called "hypotheses." In using this term he obviously did not mean to imply verbal processes in the rat but merely referred to what I have been calling cognitive maps which, it appears from his experiments, get set up in a tentative fashion to be tried out first one and then another until, if possible, one is found which works.

Finally, it is to be noted that these hypothesis experiments, like the latent learning, VTE, and "looking for the stimulus" experiments, do not, as such, throw light upon the widths of the maps which are picked up but do indicate the generally map-like and self-initiated character of learning.

For the beginning of an attack upon the problem of the width of the maps let me turn to the last group of experiments.

(5) "Spatial Orientation" Experiments

As early as 1929, Lashley reported incidentally the case of a couple of his rats who, after having learned an alley maze, pushed back the cover near the starting box, climbed out and ran directly across the top to the goal-box where they climbed down in again and ate. Other investigators have reported related findings. All such observations suggest that rats really develop wider spatial maps which include more than the mere trained-on specific paths. In the experiments now to be reported this possibility has been subjected to further examination.

In the first experiment, Tolman, Ritchie and Kalish (actually Ritchie and Kalish) used the set-up shown in Fig. 15.

This was an elevated maze. The animals ran from A across the open circular table through CD (which had alley walls) and finally to G, the food box. H was a light which shone directly down the path from G to F. After four nights, three trials per night, in which the rats learned to run directly and without hesitation from A to G, the apparatus was changed to the sun-burst shown in Fig. 16. The starting path and the table remained the same but a series of radiating paths was added.

The animals were again started at

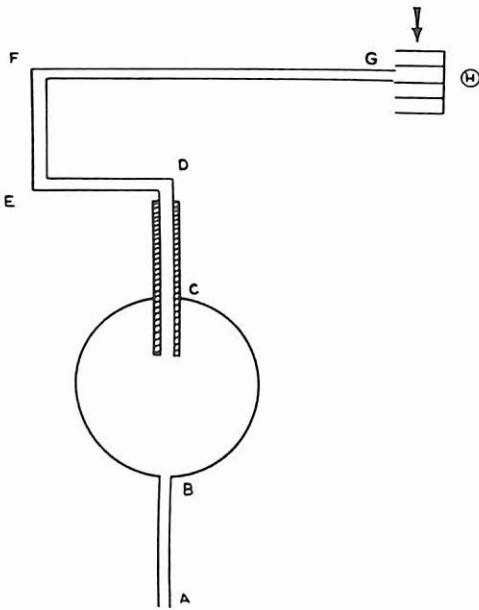


Figure 15

(From E. C. Tolman, B. F. Ritchie and D. Kalish, *Studies in spatial learning*. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 16.)

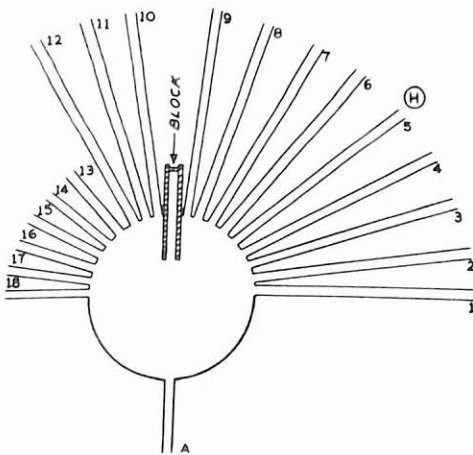


Figure 16

(From E. C. Tolman, B. F. Ritchie and D. Kalish, *Studies in spatial learning*. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 17.)

A and ran across the circular table into the alley and found themselves blocked. They then returned onto the table and began exploring practically all the radiating paths. After going out a few inches only on any one path, each rat finally chose to run all the way out on one. The percentages of rats finally choosing each of the long paths from 1 to 12 are shown in Fig. 17. It appears that there was a preponderant tendency to choose path No. 6 which ran to a point some four inches in front of where the entrance to the food-box had been. The only other path chosen with any appreciable frequency was No. 1—that is, the path which pointed perpendicularly to the food-side of the room.

These results seem to indicate that the rats in this experiment had learned not only to run rapidly down the original roundabout route but

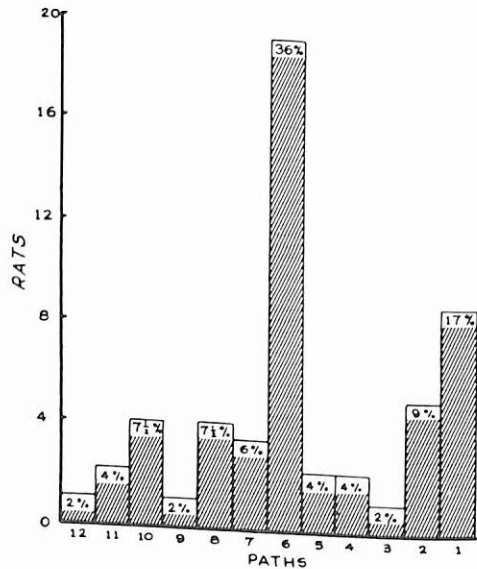


Figure 17

Number of rats which chose each of the paths (From E. C. Tolman, B. F. Ritchie and D. Kalish, *Studies in spatial learning*. I. Orientation and the short-cut. *J. exp. Psychol.*, 1946, 36, p. 19.)

also, when this was blocked and radiating paths presented, to select one pointing rather directly towards the point where the food had been or else at least to select a path running perpendicularly to the food-side of the room.

As a result of their original training, the rats had, it would seem, acquired not merely a strip-map to the effect that the original specifically trained-on path led to food but, rather, a wider comprehensive map to the effect that food was located in such and such a direction in the room.

Consider now a further experiment done by Ritchie alone. This experiment tested still further the breadth of the spatial map which is acquired. In this further experiment the rats were again run across the table—this time to the arms of a simple T. (See Fig. 18.)

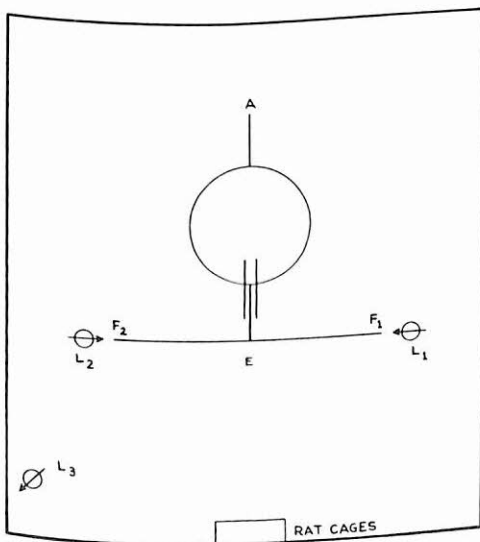


Figure 18

(From B. F. Ritchie, Ph.D. Thesis: 'Spatial learning in rats.' On deposit in the Library of the University of California, Berkeley, California.)

Twenty-five animals were trained for seven days, 20 trials in all, to find food at F_1 ; and twenty-five animals were trained to find it at F_2 . The L's in the diagram indicate lights. On the eighth day the starting path and table top were rotated through 180 degrees so that they were now in the position shown in Fig. 19. The dotted lines represent the old position. And a series of radiating paths was added. What happened? Again the rats ran across the table into the central alley. When, however, they found themselves blocked, they turned back onto the table and this time also spent many seconds touching and trying out for only a few steps practically all the paths. Finally, however, within seven minutes, 42 of the 50 rats chose one path and ran all the way out on it. The paths finally chosen by the 19 of these animals that had been fed at F_1

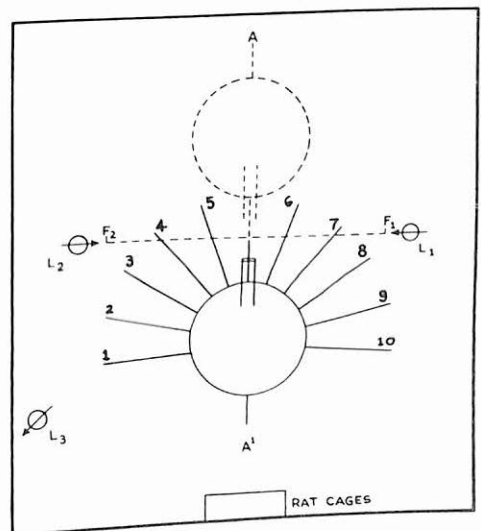


Figure 19

(From B. F. Ritchie, Ph.D. Thesis: 'Spatial learning in rats.' On deposit in the Library of the University of California, Berkeley, California.)

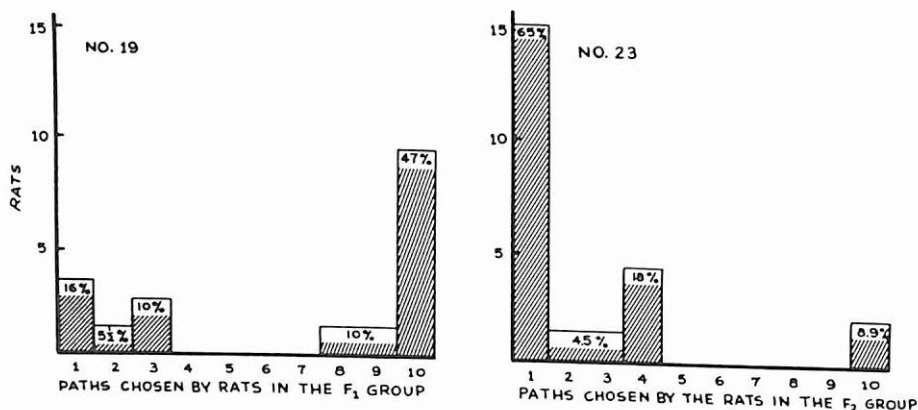


Figure 20

(From B. F. Ritchie. Ph.D. Thesis: 'Spatial learning in rats.' On deposit in the Library of the University of California, Berkeley, California.)

and by the 23 that had been fed at F₂ are shown in Fig. 20.

This time the rats tended to choose, not the paths which pointed directly to the spots where the food had been, but rather paths which ran perpendicularly to the corresponding sides of the room. The spatial maps of these rats, when the animals were started from the opposite side of the room, were thus not completely adequate to the precise goal positions but were adequate as to the correct sides of the room. The maps of these animals were, in short, not altogether strip-like and narrow.

This completes my report of experiments. There were the *latent learning experiments*, the *VTE experiments*, the *searching for the stimulus experiments*, the *hypothesis experiments*, and these last *spatial orientation experiments*.

And now, at last, I come to the humanly significant and exciting problem: namely, what are the conditions which favor narrow strip-maps and

what are those which tend to favor broad comprehensive maps not only in rats but also in men?

There is considerable evidence scattered throughout the literature bearing on this question both for rats and for men. Some of this evidence was obtained in Berkeley and some of it elsewhere. I have not time to present it in detail. I can merely summarize it by saying that narrow strip maps rather than broad comprehensive maps seem to be induced: (1) by a damaged brain, (2) by an inadequate array of environmentally presented cues, (3) by an overdose of repetitions on the original trained-on path and (4) by the presence of too strongly motivational or of too strongly frustrating conditions.

It is this fourth factor which I wish to elaborate upon briefly in my concluding remarks. For it is going to be my contention that some, at least, of the so-called 'psychological mechanisms' which the clinical psychologists and the other students of personality

have uncovered as the devils underlying many of our individual and social maladjustments can be interpreted as narrowings of our cognitive maps due to too strong motivations or to too intense frustration.

My argument will be brief, cavalier, and dogmatic. For I am not myself a clinician or a social psychologist. What I am going to say must be considered, therefore, simply as in the nature of a *rat* psychologist's *ratiocinations* offered free.

By way of illustration, let me suggest that at least the three dynamisms called, respectively, "regression," "fixation," and "displacement of aggression onto outgroups" are expressions of cognitive maps which are too narrow and which get built up in us as a result of too violent motivation or of too intense frustration.

(a) CONSIDER REGRESSION

This is the term used for those cases in which an individual, in the face of too difficult a problem, returns to earlier more childish ways of behaving. Thus, to take an example, the overprotected middle-aged woman (reported a couple of years ago in *Time Magazine*) who, after losing her husband, regressed (much to the distress of her growing daughters) into dressing in too youthful a fashion and into competing for their beaux and then finally into behaving like a child requiring continuous care, would be an illustration of regression. I would not wish you to put too much confidence in the reportorial accuracy of *Time*, but such an extreme case is not too different from many actually to be found in our mental hospitals or even sometimes in ourselves.

In all such instances my argument would be (1) that such regression results from too strong a present emotional situation and (2) that it consists in going back to too narrow an earlier map, itself due to too much frustration or motivation in early childhood. *Time's* middle-aged woman was presented by too frustrating an emotional situation at her husband's death and she regressed, I would wager, to too narrow adolescent and childhood maps since these latter had been originally excessively impressed because of overstressful experiences at the time she was growing up.

(b) CONSIDER FIXATION

Regression and fixation tend to go hand in hand. For another way of stating the fact of the undue persistence of early maps is to say that they were fixated. This has even been demonstrated in rats. If rats are too strongly motivated in their original learning, they find it very difficult to relearn when the original path is no longer correct. Also after they have relearned, if they are given an electric shock they, like *Time's* woman, tend to regress back again to choosing the earlier path.

(c) FINALLY, CONSIDER THE "DISPLACEMENT OF AGGRESSIONS ONTO OUTGROUPS"

Adherence to one's own group is an ever-present tendency among primates. It is found in chimpanzees and monkeys as strongly as in men. We primates operate in groups. And each individual in such a group tends to identify with his whole group in the sense that the group's goals become

his goals, the group's life and immortality, his life and immortality. Furthermore, each individual soon learns that, when as an individual he is frustrated, he must not take out his aggressions on the other members of his own group. He learns instead to displace his aggressions onto outgroups. Such a displacement of aggression I would claim is also a narrowing of the cognitive map. The individual comes no longer to distinguish the true locus of the cause of his frustration. The poor Southern whites, who take it out on the Negroes, are displacing their aggressions from the landlords, the southern economic system, the Northern capitalists, or wherever the true cause of their frustration may lie, onto a merely convenient outgroup. The physicists on the Faculty who criticize the humanities, or we psychologists who criticize all the other departments, or the University as a whole which criticizes the Secondary School system or, vice versa, the Secondary School system which criticizes the University—or, on a still larger and far more dangerous scene—we Americans who criticize the Russians and the Russians who criticize us, are also engaging, at least in part, in nothing more than such irrational displacements of our aggressions onto outgroups.

I do not mean to imply that there may not be some true interferences by the one group with the goals of the other and hence that the aggressions of the members of the one group against the members of the other are necessarily *wholly* and *merely* displaced aggressions. But I do assert that often and in large part they are such mere displacements.

Over and over again men are blinded by too violent motivations

and too intense frustrations into blind and unintelligent and in the end desperately dangerous hates of outsiders. And the expression of these their displaced hates ranges all the way from discrimination against minorities to world conflagrations.

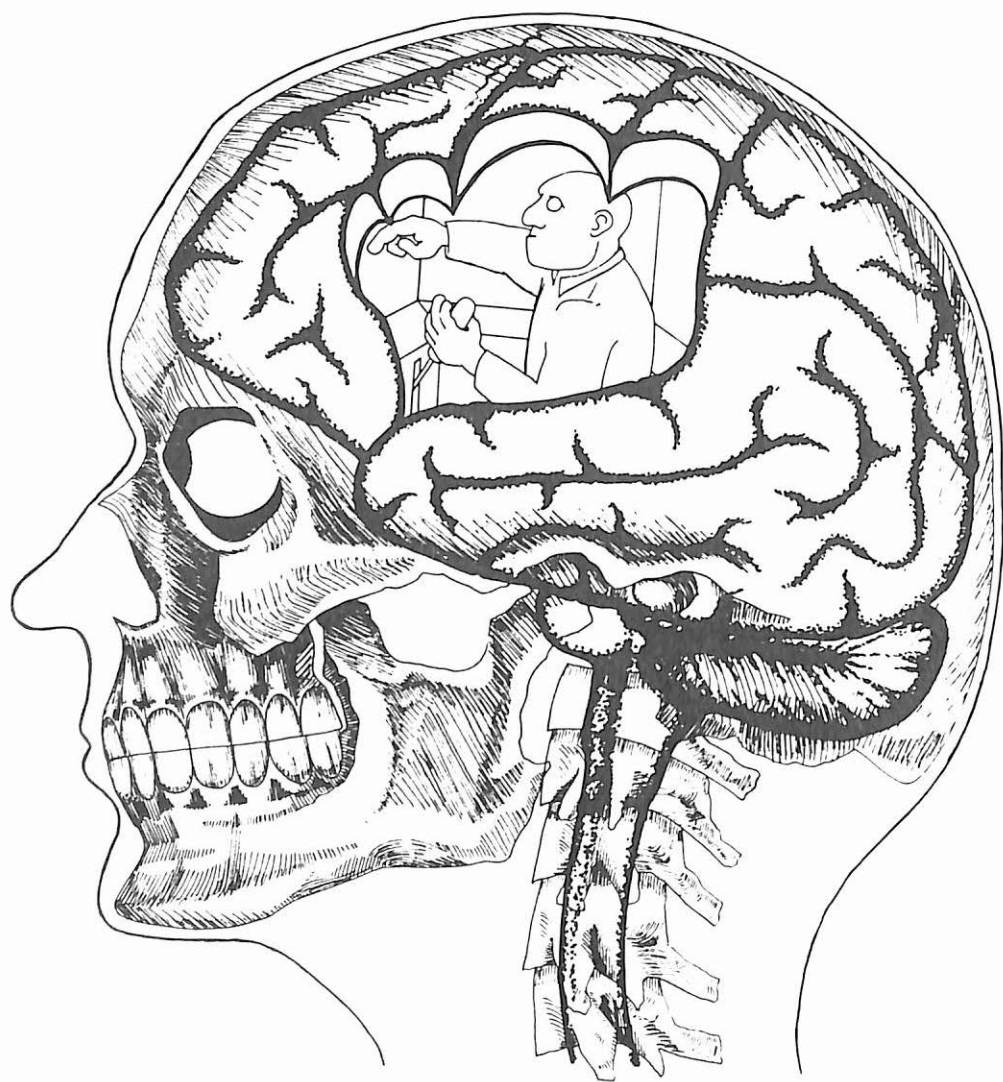
What in the name of Heaven and Psychology can we do about it? My only answer is to preach again the virtues of reason—of, that is, broad cognitive maps. And to suggest that the child-trainers and the world-planners of the future can only, if at all, bring about the presence of the required rationality (*i.e.*, comprehensive maps) if they see to it that nobody's children are too over-motivated or too frustrated. Only then can these children learn to look before and after, learn to see that there are often roundabout and safer paths to their quite proper goals—learn, that is, to realize that the well-beings of White and of Negro, of Catholic and of Protestant, of Christian and of Jew, of American and of Russian (and even of males and females) are mutually interdependent.

We dare not let ourselves or others become so over-emotional, so hungry, so ill-clad, so over-motivated that only narrow strip-maps will be developed. All of us in Europe as well as in America, in the Orient as well as in the Occident, must be made calm enough and well-fed enough to be able to develop truly comprehensive maps, or, as Freud would have put it, to be able to learn to live according to the Reality Principle rather than according to the too narrow and too immediate Pleasure Principle.

We must, in short, subject our children and ourselves (as the kindly experimenter would his rats) to the optimal conditions of moderate motivation and of an absence of unnecessary

frustrations, whenever we put them and ourselves before that great God-given maze which is our human world. I cannot predict whether or not we

will be able, or be allowed, to do this; but I *can* say that, only insofar as we *are* able and *are* allowed, have we cause for hope.



PART FOUR

biological determinants of behavior

Physiological psychologists are primarily concerned with *how* understanding what goes on inside the head will lend insight into behavioral mechanisms. As a rule the idea is not to study physiological processes per se, but to use physiological techniques and approaches to get at questions of primary interest to the psychologist. The following series of papers is particularly committed to this view. It represents some very exceptional findings.

INTRODUCTION

Physiological approaches to the study of behavior are responsible for setting many of the boundary conditions for psychological theory. While many strict behavioral problems are not interpretable in terms of physiology, physiological knowledge can save the theorist from egregious errors in his predictions. Likewise, physiological experimentation can clear the way for thinking about such issues as how the brain handles information over long periods of time. In this light, the classic work of Roger W. Sperry on the nature of the neural networks during and after embryogenesis has contributed in a major way to theories of memory and learning.

Sperry and Hibbard's article discusses an elegant experiment demonstrating that the fibers from the eye connect in a specific fashion to particular areas in the brain. Indeed, it suggests that this is not a random growth process, but a completely organized one. In brief, their results show how much of the nervous system is pre-wired and the neural circuitry pre-ordained. It has direct and compelling implications for questions dealing with the nature of intelligence, as well as other behavioral traits. Think about it.

REGULATIVE FACTORS IN THE ORDERLY GROWTH OF RETINO-TECTAL CONNEXIONS

R. W. Sperry

E. Hibbard

As a model for the study of the developmental patterning of brain pathways and connexions, the optic tract of lower vertebrates connecting the retina of the eye to the midbrain tectum offers a number of long-recognized advantages. Much of our current understanding of the developmental organization of brain circuits stems from experiments on this fairly simple and accessible optic system. There is good reason to think that the basic developmental mechanisms found to operate in the prefunctional shaping of this system for visual behaviour have wide application to the developmental patterning of the nervous system in general.

Strange as it may seem today, the first experiments on the growth of behavioural organization in this system were designed to determine whether the basic visual properties, like perception of directionality, size, movement, contour, and so on, are installed directly by the growth process itself or

have to be organized through use and behavioural adjustment (Sperry, 1965). This was a little over 25 years ago now, when it was still believed that the growth of nerve fibres was non-selective, subject only to mechanical guidance. Chemical and electrical guidance appeared to have been ruled out, and several lines of evidence seemed to show that selective synaptic connexions are unimportant, at least for orderly behaviour. "Instinct" was still a bad word in science and the whole concept of the inheritance of behavioural patterns remained anathema to prevailing doctrines in psychology. No developmental machinery was then known or available, even in theory, by which the highly ordered and precisely designed neural connexions for behaviour could be "grown into" the nervous system without the aid of function. Neurobiotaxis, with its dependence on function for selectivity, was still the prevailing favourite among explanatory concepts. Even the old master, Cajal, had been willing to leave it to function when it came to the detailed central adjustments of synaptic associations for behaviour.

We have come a long way since

Reprinted from the Ciba Foundation Symposium on Growth of the Nervous System, edited by G. E. W. Wolstenhove and M. O'Connor, J. and A. Churchill, Ltd., London, 1968, pp. 41-52. Reproduced by permission.

then, of course, and the issue of growth versus function seems now to be quite settled and dead, at least in the case of retino-tectal connexions. The question is nevertheless of some current concern with respect to many of the other fibre tracts of the brain, particularly in man. It would seem fairly safe to speculate in this regard that the patterning of most of the long fibre systems of the CNS, even in man, is primarily a problem of developmental mechanics, not one of learning. Another way of saying the same thing is that, if learning involves changes in the morphology of nerve connexions at all, these would appear to be confined to dendritic relations and local neuron circuits without affecting very much the patterning of the long fibre systems of the brain.

Once the growth-versus-function issue was settled in favour of growth (Sperry, 1944), attention was turned to the nature of possible developmental mechanisms. The next general question—and one still pertinent and basic to the nervous system as a whole—concerned the extent to which orderly selective function is dependent on selectivity and orderliness in the underlying neural connexions. As late as 1954 it was being suggested that orderly visual recovery might be explained even though the divided optic fibres reconnected in a random diffuse pattern, with retino-tectal communication based on a physiological coding-decoding scheme. Similarly the pre-1940 Resonance Principle of Weiss (1937) and a more recent interpretation along similar lines proposed by Szentágothai (1961) and Székely (1966) would provide for selective communication regardless of absence of selectivity in the morphology of the fibre connexions.

Evidence that the normal topo-

graphical projection on the tectum is in fact restored after section and scrambling of the optic fibres came first from experiments in which localized lesions were made in the frog tectum. The position and size of the resultant scotoma or blind area were found to be the same whether the localized lesion was made before division of the optic nerve or after regeneration (Sperry, 1944)—a result obtainable only if the regenerated projection map had been restored in its original pattern. Further confirmation on a more refined scale has since come from electrical mapping methods, particularly in the studies of Gaze and his associates (reviewed in Gaze, 1967). Confirmation has also been obtained with histological procedures (Attardi and Sperry, 1960, 1963). The patterning of the optic fibre pathways and their terminal connexions was shown in the latter studies to be selectively predetermined according to the locus of origin of the optic fibre in the retinal field. Fibres from different retinal sectors were found to exhibit different growth patterns, like different species of plants, by which they selectively regained their various predetermined target zones in the tectal field.

Other experiments, with H. L. Arora and A. J. Limpo (see Sperry, 1965), have shown that colour perception—like position, directionality, movement, size, and contour perception—undergoes an orderly reinstatement in optic nerve regeneration. In summary, all the basic functional attributes of vision appear thus to be restored in regeneration of the scrambled optic fibres. Endogenous physiological properties of different sensory and neural cell types were presumed to be involved, as well as synaptic connexions. Where the experiments have been extended into prefunctional stages in-

volving initial nerve outgrowth, similar results have been obtained (see Székely, 1966).

As early as 1942 it was inferred (Sperry, 1942) that optic fibres arising from different retinal loci must be distinguished from one another according to the location of their ganglion cell bodies in the retinal field, "probably through differential physical-chemical properties induced in them by differentiation of the optic cup in development". It was suggested further (Sperry, 1945) that the retina must undergo a gradient or field type of differentiation along at least two separate axes and that contralateral eye transplants might demonstrate that the anteroposterior gradient is established before the dorsoventral gradient. A corresponding biochemical specification of the tectal field was likewise inferred and it was postulated that differential affinities must exist between the tectal neurons and the optic fibres. With supporting results from the vestibular, cutaneous, central and motor systems (Sperry, 1951*a*, *b*) the basic concepts were extended to the development of the nervous system generally. This material has been reviewed in some detail on four separate occasions in the past three years (Gaze, 1967; Jacobson, 1966; Sperry, 1965; Székely, 1966) and does not need further amplification.

Recent work on the optic system has been aimed at more detailed analysis of the various phases of the postulated specification processes, chemical-affinity mechanisms, and related factors. Gaze, Jacobson and Székely (1963) tested the elasticity of the tectal gradients in response to ingrowth of fibres from compound eyes composed of two nasal or two temporal half-fields. In this interesting experiment it appeared

as if the recipient tectal half-field had spread to cover the entire tectum, producing a rather radical change in the pattern of tectal differentiation. However, as we suggested to Dr. Székely in a seminar at the Massachusetts Institute of Technology a few years ago (see also Sperry, 1965), one should probably reserve judgment on the interpretation of these results until the effects of tectal hyperplasia and hypoplasia, in response to the overload and diminished innervation, can be ruled out. The spread of the innervated half of the tectal gradient into the denervated half may be more apparent than real and more mechanical than chemical. Related problems are mainly of embryological interest. They bear primarily on the developmental dynamics of the morphogenetic field and have only remote significance for problems of brain function.

A question with more direct neurological implications concerns the degree of local resolution and overlap in the optic fibre projection on the tectum. How large is the terminal arborization field of a single fibre and with how many tectal cells over how large a dendritic field does one optic fibre synapse? How much overlap is there in the synaptic zone of neighbouring individual fibres? How do these compare in the normal and in the regenerated condition? The answers to these and related questions may be expected to differ in different species and for different parts of the retina. At this stage a rough approximation would be helpful in judging the higher degrees of precision required in the regeneration process to restore normal vision. The maps based on electrical procedures are difficult to interpret in this connexion because of the variables associated with changes in amplification and the uncertainty

whether the evoked potentials come from optic fibre terminals or from post-synaptic firing in the tectal neurons.

Using Golgi staining of the normal two-day-old chick tectum we have seen terminal arborizations of the optic fibres in considerable numbers, measuring about $75\ \mu$ or approximately 1/200th of the total arc of the central tectal field. A somewhat lower order of refinement is probably present in the cichlid fishes like the angel fish *Pterophyllum scalare* and *Astronotus ocellatus*, in which we have seen visual recovery with little if any loss in visual acuity. A large—22 per cent—loss of acuity on the average was reported in another study on *A. ocellatus* by Weiler (1966). The extent to which the regenerative process is able to reproduce the normal degree of resolution in visual function needs further study. It is difficult to judge the degree of selective precision of the growth of optic fibres from the pattern established normally because of the complex sequence of possible sustaining factors present in normal development. Observations on the interposition of obstacles and other types of interruptions in the fibre pattern within the tectum (see below) indicate that considerable correction is possible in regrowth within the plexiform layer close to the target zone.

Some years ago one of us (E. H.) noted a very striking similarity in the layered differentiation of retina and tectum in the developing frog tadpole. This is illustrated in Fig. 1 for the ten-week-old tadpole of *Rana pipiens*. While the character of the layering and the thickness of the retina remains fairly uniform throughout, that of the tectum varies somewhat from the dorsomedial to the ventrolateral portions of the lobe, primarily in the thickness of the periventricular cell layer. The

segment shown in Fig. 1 represents the central part of the lobe, corresponding approximately to the position where the centre of the visual field would be projected. The photographs are shown at identical magnifications, the distance from the innermost to the outermost uniform cell layers being $175\ \mu$ in both cases. The relative position of the optic nerve fibre layer is the same in both retina and tectum, but the thickness of the tectum is somewhat greater because of the marginal layer overlying the optic fibres. Although the number of cells in each layer arranged perpendicular to the surface is approximately the same (one in the ganglion cell and superficial grey matter, eight to ten in the inner nuclear and central grey matter, two in the outer nuclear and periventricular grey matter), the surface area of the tectum is less than 60 per cent of that of the retina. A considerable amount of convergence between retinal ganglion cells and terminal tectal sites seems indicated. The numbers and positions of cell layers, the distribution and thickness of plexiform layers, and the relative position of outgoing and incoming optic fibres in these structures appear to be more than coincidental.

We have been investigating in a preliminary way, with Dr. Arora (reviewed in Sperry, 1965), the forces that control the growth of optic fibres in various parts of the afferent course into the tectum, using surgical deflections, interruptions and transplantations. When the medial and lateral brachia of the optic tract are cut short and crossed at the anterior tectal pole the bulk of the sprouting fibres promptly turn back to regain their normal channel. When the lateral tract, cut long, is transposed deep into the medial pathway, most fibres enter the medial tract and follow it on

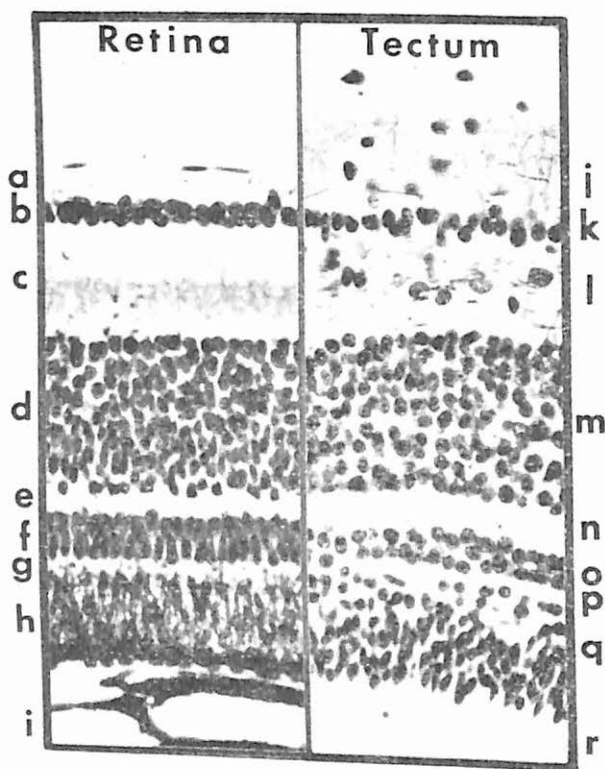


Figure 1
Comparison of retinal and tectal stratification in a ten-week-old tadpole of *Rana pipiens*. $\times 400$

- (a) optic fibre layer
- (b) ganglion cell layer
- (c) inner plexiform layer
- (d) inner nuclear layer
- (e) outer plexiform layer
- (f) outer nuclear layer
- (g) limitans
- (h) photoreceptor and pigment cell layers
- (i) choroid vessels

- (j) marginal and optic layers
- (k) superficial grey layer
- (l) deep medullary layer
- (m) central grey layer
- (n) deep white layer
- (o) periventricular grey layer
- (p) periventricular fibrous layer
- (q) subependymal granular and ependymal layer
- (r) ventricle

across the equator of the tectum into the ventral quadrants before they enter the plexiform layer to synapse in their original zone. When contralateral fibres are superimposed on a fully innervated optic tract they form a

plexiform zone only in the corresponding mirror locus of the opposite tectum.

The ability of the optic fibres to find their way to their proper tectal stations even from grossly abnormal

starting positions has been investigated by surgical deflection of the cut optic nerve. The central end of the divided optic nerve was apposed to the root of the oculomotor nerve in such a way that the regenerating optic fibres then entered the brain below and caudal to the tectum, near the point where the oculomotor fibres leave. Gaze (1959) had previously reported a case in which the optic nerve inadvertently entered the brain along this route. He was able to record a few evoked potentials on the contralateral tectum corresponding to points of stimulation on the retina from which the nerve had originated, but was unable to elicit optokinetic responses through this nerve. In the present series, the optic nerve was transposed caudally to a point near the roots of the oculomotor and trigeminal nerves in 14 *Xenopus laevis* tadpoles. In three cases, both optic nerves entered the brain through the oculomotor root, and most of the optic fibres succeeded in making their way back to the tectum. However, these connexions were made with the ipsilateral tectum—not by following the oculomotor fibres back towards the nucleus of cranial nerve III as might have been expected, but rather by decussating twice, once at the level of the IIIrd root and again in the region of the optic chiasma (see Fig. 2). Behaviourally, this led to reversed optokinetic responses in both directions in two animals, and spontaneous circling to the right in the third (Hibbard, 1967).

In two other cases, the left optic nerve regenerated over the surface of the brain above the meninges. In one of these latter, the aberrant optic nerve was cut again above the mid-brain and poked down through a hole into the dorsal surface of the ipsi-

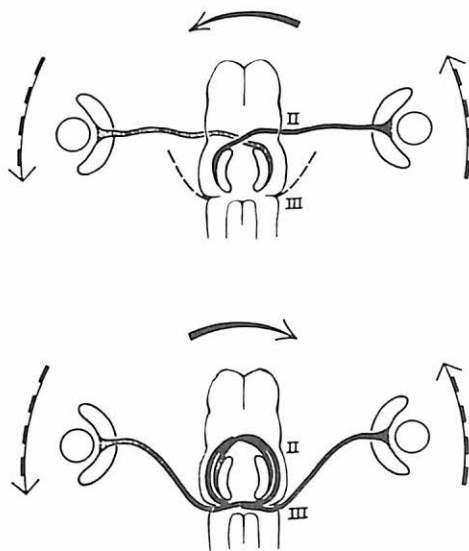


Figure 2

Diagrams illustrating the courses taken by optic fibres from the eye to the tectum in normal *Xenopus* tadpole (upper), and tadpole in which the optic nerves have been diverted into the oculomotor nerve roots (lower). The response of the normal animal to counter-clockwise rotation of an optokinetic drum, indicated by banded arrows, is to turn to the left. The response of the experimental animal, in which optic fibres cross twice and terminate in the ipsilateral tectal lobes, is to turn in the opposite direction.

lateral tectum. Thirteen days later this animal began to show spontaneous circling to the left at a rate of about 12 turns per minute. The deflected optic nerves in seven other cases strayed caudally and entered the brain along the trigeminal nerve root. In traversing the ophthalmic ramus, ganglion, and root of the Vth nerve the optic fibres in these cases remained fasciculated and did not intermingle with trigeminal fibres. The optic fibres

could not be followed any distance into the brain within the trigeminal tract and there was no evidence of visual recovery under these conditions.

The orderly, near-parallel distribution of optic fibres in the "parallel" or "radiation" layer of the tectum has yet to be explained. What are the invisible guidelines or forces that keep the fibres from crossing and interspersing and from curling up or dipping down? Are the invisible guidelines primarily mechanical or chemical or

a combination of both? In an effort to get some answers to such questions we have made various kinds of experimental interruptions and derangements of the parallel layer in adult fishes, some examples of which are shown in Fig. 3. This work is still in progress, but a few points seem to be emerging. Although the deflected fibres seem not to be able to effect much change of course within the parallel layer itself, the same fibres may show extensive readjustment in the plexi-

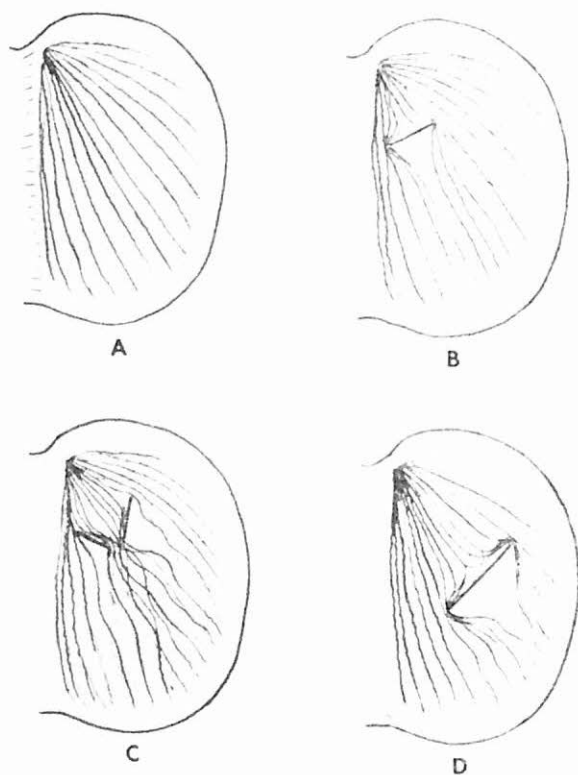


Figure 3
Routes taken by regenerating optic fibres from the medial tract in the tectum when blocked by tantalum foil barriers. (A) Normal. (B) Small obstacle with fibre-free shadow behind it. (C) Two barriers through which fibres are funneled. (D) Large barrier with fine fibres in the shadow area behind it.

form layer just below. This would appear to permit a considerable degree of target correction near the terminations and to reduce the degree of precision required at choice points in more distal portions of the optic fibre channels. The deflected fascicles undulate through the plexiform layer in smooth curves and gradually straighten out. We have not been able to substantiate Leghissa's (1955) description of a criss-cross overlap among the optic fibres of the parallel layer at the posterior equator.

If, after the optic fibres in the parallel fascicles have been cut, barriers of tantalum foil are inserted across their course, the regenerating fibres pass around the block in a smooth flow, leaving a "shadow" area in the parallel layer relatively clear of fibres immediately behind the block. If two strips of foil are placed at angles of 45° to the parallel bundles in such a way as to form a funnel through which the regenerating fibres will pass, the deflected fibres pass through the narrow opening, at angles of 90° and less, later crossing and re-crossing other fibres largely below or sometimes above the parallel layer proper. They

gradually refill the parallel fascicles distal to the block, but whether the fibres are regaining the same channels as they would normally occupy or are merely refilling any empty pathways is not clear. The undulating pattern suggests an inertial component in the fibre growth, with oscillations and corrections of deviations from course by negative feedback.

Correction of the course of the outgrowing axon, apparently in response to an anteroposterior gradient effect, has been clearly shown in abnormally positioned Mauthner cell axons (Hibbard, 1965). If the segment of hindbrain containing the Mauthner cell is reversed at an embryonic stage before the outgrowth of the axon, the axon first grows forward, towards the original posterior end of the grafted segment, but then re-curves to proceed down the length of the spinal cord. In embryos joined back-to-back near the end of the hindbrain (see Fig. 4) the Mauthner axon proceeds along its typical course into the twin, but travelling in the caudo-rostral direction directly opposite to normal. Presumably the microfilament flare at the growing tip does not span sufficient rostro-



Figure 4

Two-headed *Xenopus* embryo formed by joining anterior halves of two embryos to study the effect on the course of Mauthner's axons.

caudal differences to overcome the inertial growth component (Swisher and Hibbard, 1967). The direction of advance of a single growing fibre is probably better viewed as a reflection of the algebraic sum of the combined adhesive forces of all the active micro-filaments than as the sequence of single strong filament successes. Van der Loos (1965), studying abnormally oriented pyramidal cells in the cortex, has suggested that the initial outgrowth of the axon, in some instances at least, as well as the spread of the dendrites, is intrinsically determined within the cell and is dependent upon the position of the cell, while further outgrowth of the axon is determined by extrinsic factors.

Summary

The nerve tract connecting eye to brain, together with its afferent and efferent association within the retina at one end and within the midbrain tectum at the other, may be regarded as a model system for the study of the developmental organization of brain pathways and synaptic connexions. The physiological and behavioural properties of this system indicate a high degree of complexity, precision, diversity, and refinement in the underlying neural relations; any understanding of the regulative factors by which this optic system is organized in growth should therefore take us a long way towards a general understanding of the basic developmental mechanisms for the inherent patterning of brain pathways and connexions.

The orderly formation of retino-tectal connexions, both in development and in regeneration, has been studied extensively in experiments involving surgical, behavioural, histological, electrical, biochemical, and re-

lated methods, applied separately and in combination. The principal regulative factors disclosed so far by the accumulated evidence have been outlined and some of the behavioural and histological results briefly illustrated. Some newer observations have been presented that deal with the patterning of fibre routes and connexions within the tectum, with a correspondence in retinal and tectal stratification, with the growth patterns of transposed optic nerves, and with the oriented advance of individual axons. An attempt is made to show how the developmental mechanisms inferred from experiments on the retino-tectal relations of lower vertebrates could also account in principle for the ontogenetic organization of the much more complicated mammalian visual system.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institute of Mental Health, MH3372, and by the Hixon Fund of the California Institute of Technology. The technical assistance of Ruth T. Johnson is gratefully acknowledged.

References

- ATTARDI, D. G., and SPERRY, R. W. (1960). *Physiologist, Wash.*, 3, 12.
- ATTARDI, D. G., and SPERRY, R. W. (1963). *Expl. Neurol.*, 7, 46-64.
- GAZE, R. M. (1959). *J. Physiol., Lond.*, 146, 40.
- GAZE, R. M. (1967). *A. Rev. Physiol.*, 29, 59.
- GAZE, R. M., JACOBSON, M., and SZÉKELY, G. (1963). *J. Physiol., Lond.*, 165, 484-499.
- HIBBARD, E. (1965). *Expl. Neurol.*, 13, 289-301.

- HIBBARD, E. (1967). *Expl. Neurol.*, **19**, 350-356.
- JACOBSON, M. (1966). In *Current Status of some Major Problems in Developmental Biology* (25th Symp. Soc. Devel. Biol.), pp. 339-383, ed. Locke, M. New York: Academic Press.
- LEGHISSA, S. (1955). *Z. Anat. EntwGesch.*, **118**, 427-463.
- SPERRY, R. W. (1942). *Anat. Rec.*, **84**, 470.
- SPERRY, R. W. (1944). *J. Neurophysiol.*, **7**, 57-69.
- SPERRY, R. W. (1945). *J. Neurophysiol.*, **8**, 15-28.
- SPERRY, R. W. (1951a). In *Handbook of Experimental Psychology*, pp. 236-280, ed. Stevens, S. S. New York: Wiley.
- SPERRY, R. W. (1951b). *Growth* **15**, Suppl., 63-87.
- SPERRY, R. W. (1965). In *Organogenesis*, ch. 6, ed. De Haan, R. L., and Ursprung, H. New York: Holt, Rinehart & Winston.
- SWISHER, J. E., and HIBBARD, E. (1967). *J. exp. Zool.*, **165**, 433-440.
- SZÉKELY, G. (1966). *Adv. Morphoghen.*, **5**, 181.
- SZENTÁGOTHAÍ, J. (1961). In *Brain and Behavior*, vol. 1, ed. Brazier, M. A. B. Washington, D.C.: American Institute of Biological Sciences.
- VAN DER LOOS, H. (1965). *Bull. Johns Hopkins Hosp.*, **117**, 228-250.
- WEILER, I. J. (1966). *Expl. Neurol.*, **15**, 377.
- WEISS, P. A. (1937). *J. comp. Neurol.*, **66**, 481-535.

INTRODUCTION

What is it about the brain that allows us to differentiate squares from triangles, curves from straight lines, or more generally any figure from the background? In other words, how does the visual system work? For years psychologists and physiologists have puzzled over this, and only recently has the visual system been the subject of direct physiological analysis. In one of their outstanding papers, Hubel and Wiesel demonstrate how a sensory system of the brain can be systematically analyzed and the physiological properties of the system delineated in a way that gives us hints on the nature of visual perception. Using tiny micro-electrodes, physiological responses from single neurons are studied while vision patterns of assorted nature are presented in full view of a cat. While recording from one neuron in the visual cortex they are able to show how that neuron responds to a figure presented in the external world. From such studies they build an elaborate and intricate, but highly fascinating, story on the physiological basis of visual perception.

RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX

D. H. Hubel
T. N. Wiesel

In the central nervous system the visual pathway from retina to striate cortex provides an opportunity to observe and compare single unit responses at several distinct levels. Patterns of light stimuli most effective in influencing units at one level may no longer be the most effective at the next. From differences in responses at successive stages in the pathway one may hope to gain some understanding of the part each stage plays in visual perception.

By shining small spots of light on the light-adapted cat retina Kuffler (1953) showed that ganglion cells have concentric receptive fields, with an 'on' centre and an 'off' periphery, or vice versa. The 'on' and 'off' areas within a receptive field were found to be mutually antagonistic, and a spot restricted to the centre of the field was more effective than one covering the whole receptive field (Barlow, Fitz-Hugh & Kuffler, 1957). In the freely moving light-adapted cat it was found that the great majority of cortical cells studied gave little or no response to light stimuli covering most of the animal's visual field, whereas small spots

shone in a restricted retinal region often evoked brisk responses (Hubel, 1959). A moving spot of light often produced stronger responses than a stationary one, and sometimes a moving spot gave more activation for one direction than for the opposite.

The present investigation, made in acute preparations, includes a study of receptive fields of cells in the cat's striate cortex. Receptive fields of the cells considered in this paper were divided into separate excitatory and inhibitory ('on' and 'off') areas. In this respect they resembled retinal ganglion-cell receptive fields. However, the shape and arrangement of excitatory and inhibitory areas differed strikingly from the concentric pattern found in retinal ganglion cells. An attempt was made to correlate responses to moving stimuli with receptive field arrangements. Some cells could be activated from either eye, and in these binocular interaction was studied.

Methods

In this series of experiments twenty-four cats were used. Animals were anaesthetized with intraperitoneal thiopental sodium (40 mg/kg) and light

anaesthesia was maintained throughout the experiment by additional intraperitoneal injections. The eyes were immobilized by continuous intravenous injection of succinylcholine; the employment of this muscle relaxant made it necessary to use artificial respiration. Pupils of both eyes were dilated and accommodation was relaxed by means of 1% atropine. Contact lenses used with a suitably buffered solution prevented the corneal surfaces from drying and becoming cloudy. The lids were held apart by simple wire clips.

A multibeam ophthalmoscope designed by Talbot & Kuffler (1952) was used for stimulation and viewing the retina of the left eye. Background illumination was usually about 0.17 log. metre candles (m.c.), and the strongest available stimulus was 1.65 log. m.c. Many sizes and shapes of spots of light could be produced, and these were well focused on the retina. Stimulus durations were of the order of 1 sec.

For binocular studies a different method of light stimulation was used. The animal faced a large screen covering most of the visual field. On this screen light spots of various sizes and shapes were projected. The light source was a tungsten filament projector mounted on an adjustable tripod. Stimuli could be moved across the screen in various directions and with different speeds. Spots subtending an angle as small as 12 min of arc at the cat's eyes could be obtained, but generally 0.5–1° spots were used for mapping receptive fields. (Dimensions of stimuli are given in terms of equivalent external angles; in the cat 1 mm on the retina subtends about 4°.) Spots were focused on the two retinas with lenses mounted in front of the cat's eyes. Lenses for focusing were selected by using a retinoscope. Spot intensi-

ties ranged from -0.76 to 0.69 log. cd/m². A background illuminance of -1.9 log. cd/m² was given by a tungsten bulb which illuminated the whole screen diffusely. Intensities were measured by a Macbeth Illuminometer. Values of retinal illumination corresponding to these intensities (Talbot & Kuffler, 1952, Fig. 4) were within the photopic range but were lower than those employed with the ophthalmoscope. Whenever the two methods of stimulation were checked against each other while recording from the same unit they were found to give similar results. This principle of projecting light spots on a screen was described by Talbot & Marshall (1941). Areas responsive to light were marked on sheets of paper fixed on the screen, in such a way as to indicate whether the responses were excitatory or inhibitory. The sheets of paper then provided permanent records of these responses, and showed the shape, size, and orientation of the regions.

Single unit activity was recorded extracellularly by techniques described previously (Hubel, 1959). A hydraulic micro-electrode positioner was attached to the animal's skull by a rigidly implanted plastic peg. The cortical surface was closed off from the atmosphere to minimize respiratory and vascular movements of the cortex (Davies, 1956). This method gave the stability needed for thorough exploration of each receptive field, which often took many hours. Electrodes were electrolytically sharpened tungsten wires insulated with a vinyl lacquer (Hubel, 1957). Cathode follower input and a condenser-coupled pre-amplifier were used in a conventional recording system.

Recordings were made from parts of the lateral gyrus extending from its posterior limit to about Horsley-

Clarke frontal plane 10. At the end of each penetration an electrolytic lesion was made (Hubel, 1959) and at the end of the experiment the animal was perfused, first with normal saline and then with 10% formalin. The borders of the trephine hole were marked with India ink dots and the brain was removed from the skull and photographed. Paraffin serial sections were made in the region of penetration and stained with cresyl violet. These sections showed that all units described were located in the grey matter of the striate cortex. Correlation between location of units in the striate cortex and physiological findings will not be dealt with in this paper.

There is evidence that cortical cells and afferent fibres differ in their firing patterns and in their responses to diffuse light (Hubel, 1960). The assumption that the spikes recorded were from cell bodies is based on these differences, as well as on electrophysiologic criteria for distinguishing cell-body and fibre spikes (Frank & Fuortes, 1955; Hubel, 1960).

Results

Several hundred units were recorded in the cat's striate cortex. The findings to be described are based on thorough studies of forty-five of these, each of which was observed for a period of from 2 to 9 hrs. Times of this order were usually required for adequate analysis of these units.

In agreement with previous findings in the freely moving light-adapted cat (Hubel, 1959) single cortical units showed impulse activity in the absence of changes in retinal illumination. Maintained activity was generally less than in freely moving animals, and ranged from about 0.1–10 im-

pulses/sec. The low rate was possibly due to light barbiturate anaesthesia, since on a number of occasions deepening the anaesthesia resulted in a decrease of maintained activity. This need not mean that all cortical cells are active in the absence of light stimuli, since many quiescent units may have gone unnoticed.

In most units it was possible to find a restricted area in the retina from which firing could be influenced by light. This area was called the receptive field of the cortical unit, applying the concept introduced by Hartline (1938) for retinal ganglion cells. The procedure for mapping out a receptive field is illustrated in Fig. 1. Shining a 1° spot (250μ on the retina) in some areas of the contralateral eye produced a decrease in the maintained activity, with a burst of impulses when the light was turned off (Fig. 1*a, b, d*). Other areas when illuminated produced an increase in firing (Fig. 1*c, e*). The complete map, illustrated to the right of the figure, consisted of a long, narrow, vertically oriented region from which 'off' responses were obtained (triangles), flanked on either side by areas which gave 'on' responses (crosses). The entire field covered an area subtending about 4° . The elongated 'off' region had a width of 1° and was 4° long.

Most receptive fields could be subdivided into excitatory and inhibitory regions. An area was termed excitatory if illumination produced an increase in frequency of firing. It was termed inhibitory if light stimulation suppressed maintained activity and was followed by an 'off' discharge, or if either suppression of firing or an 'off' discharge occurred alone. In many units the rate of maintained activity was too slow or irregular to demonstrate inhibition during illumination,

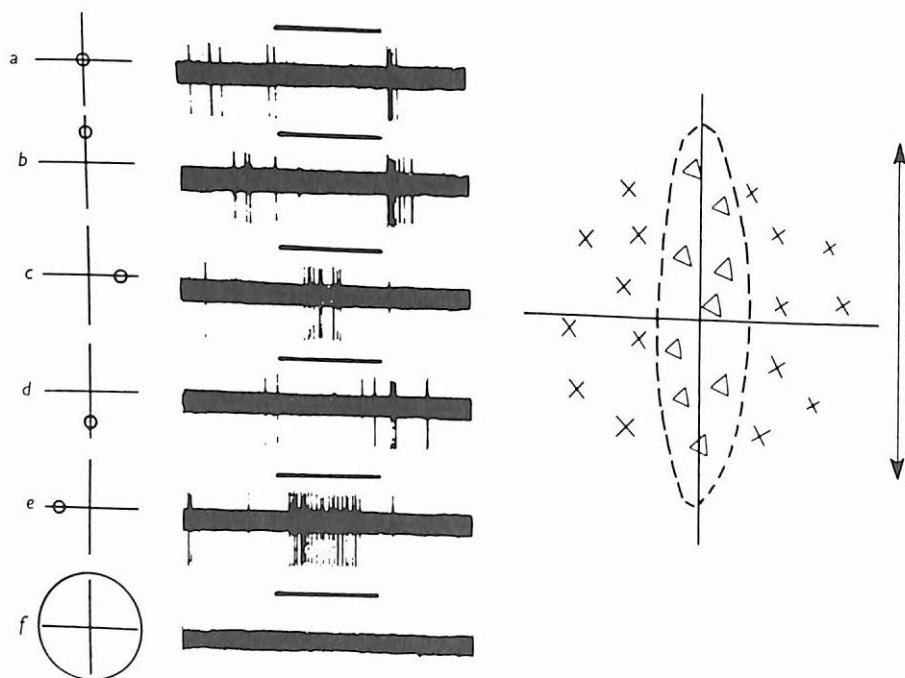


Figure 1

Responses of a cell in the cat's striate cortex to a 1° spot of light. Receptive field located in the eye contralateral to the hemisphere from which the unit was recorded, close to and below the area centralis, just nasal to the horizontal meridian. No response evoked from the ipsilateral eye. The complete map of the receptive field is shown to the right. $\times$, areas giving excitation; Δ , areas giving inhibitory effects. Scale, 4° . Axes of this diagram are reproduced on left of each record, *a*, 1° (0.25 mm) spot shone in the centre of the field; *b-e*, 1° spot shone on four points equidistant from centre; *f*, 5° spot covering the entire field. Background illumination 0.17 log. m.c. Stimulus intensity 1.65 log. m.c. Duration of each stimulus 1 sec. Positive deflexions upward.

and only an 'off' discharge was seen. It was, however, always possible to demonstrate inhibitory effects if the firing rate was first increased by stimulation of excitatory regions.

As used here, 'excitatory' and 'inhibitory' are arbitrary terms, since both inhibition and excitation could generally be demonstrated from both regions, either during the light stimulus or following it. We have chosen to denote receptive field regions according to effects seen during the stim-

ulus. Furthermore, the word 'inhibition' is used descriptively, and need not imply a direct inhibitory effect of synaptic endings on the cell observed, since the suppression of firing observed could also be due to a decrease in maintained synaptic excitation.

When excitatory and inhibitory regions (used in the sense defined) were stimulated simultaneously they interacted in a mutually antagonistic manner, giving a weaker response than when either region was illuminated

alone. In most fields a stationary spot large enough to include the whole receptive field was entirely without effect (Fig. 1*f*). Whenever a large spot failed to evoke responses, diffuse light stimulation of the entire retina at these intensities and stimulus durations was also ineffective.

In the unit of Fig. 1 the strongest inhibitory responses were obtained with a vertical slit-shaped spot of light covering the central area. The greatest 'on' responses accompanied a stimulus confined to the two flanking regions. Summation always occurred within an area of the same type, and the strongest response was obtained with a stimulus having the approximate shape of this area.

In the unit of Figs. 2 and 3 there was weak excitation in response to a circular 1° spot in the central region. A weak 'off' response followed stimulation in one of the flanking areas (Fig. 2*a, b*). There was no response to an 8° spot covering the entire receptive

field (Fig. 2*c*). The same unit was strongly activated by a narrow slit-shaped stimulus, measuring 1° by 8° , oriented vertically over the excitatory region (Fig. 3*A*). In contrast, a horizontal slit of light was completely ineffective, despite the fact that the central area was capable of evoking a response when stimulated alone (Fig. 2*a*). As the optimum (vertical) orientation of the slit was approached, responses appeared and rapidly increased to a maximum.

These findings can be readily understood in terms of interacting excitatory and inhibitory areas. The strength of the response to a vertically oriented slit is explained by summation over the excitatory region and by the exclusion of inhibitory regions. When parts of the inhibitory flanking areas were included by rotating the slit, responses were reduced or abolished. Thus a horizontal slit was ineffective because it stimulated a small portion of the central excitatory area,

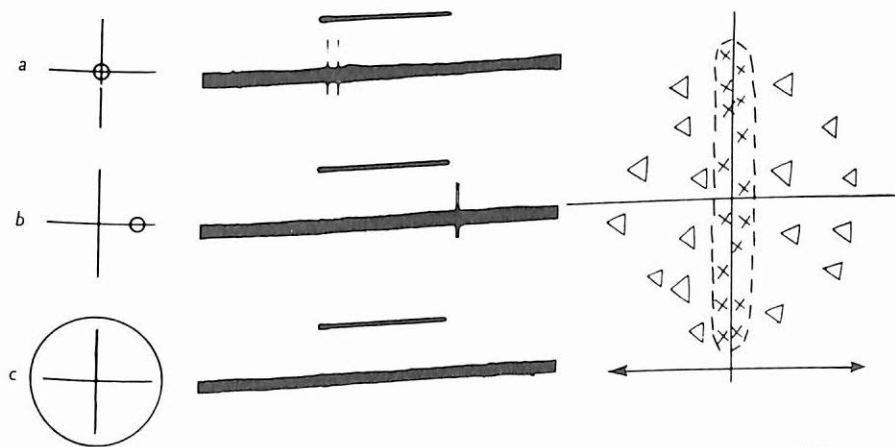
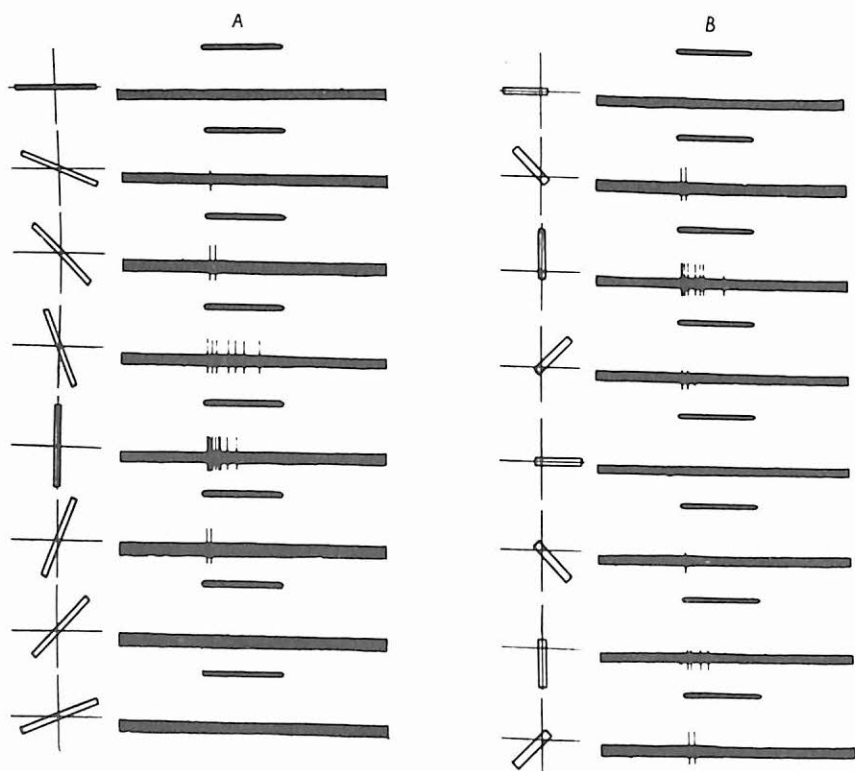


Figure 2

Responses of a unit to stimulation with circular spots of light. Receptive field located in area centralis of contralateral eye. (This unit could also be activated by the ipsilateral eye.) *a*, 1° spot in the centre region; *b*, same spot displaced 3° to the right; *c*, 8° spot covering entire receptive field. Stimulus and background intensities and conventions as in Figure 1. Scale, 6° .



Same unit as in Figure 2. *A*, responses to shining a rectangular light spot, $1^\circ \times 8^\circ$; centre of slit superimposed on centre of receptive field; successive stimuli rotated clockwise, as shown to left of figure *B*, responses to a $1^\circ \times 5^\circ$ slit oriented in various directions, with one end always covering the centre of the receptive field: note that this central region evoked responses when stimulated alone (Figure 2*a*). Stimulus and background intensities as in Figure 1; stimulus duration 1 sec.

and larger portions of the antagonistic regions.

Some units were not responsive enough to permit mapping of receptive fields with small light spots. In these the effective stimulus pattern could be found by changing the size, shape and orientation of the stimulus until a clear response was evoked. Often when a region with excitatory or inhibitory responses was established the neighbouring opposing areas in the receptive field could only be demonstrated indirectly. Such an indirect

method is illustrated in Fig. 3*B*, where two flanking areas are indicated by using a short slit in various positions like the hand of a clock, always including the very centre of the field. The findings thus agree qualitatively with those obtained with a small spot (Fig. 2*a*).

Receptive fields having a central area and opposing flanks represented a common pattern, but several variations were seen. Some fields had long narrow central regions with extensive flanking areas (Figs. 1-3); others had

a large central area and concentrated slit-shaped flanks (Figs. 6, 9, 10). In many fields the two flanking regions were asymmetrical, differing in size and shape; in these a given spot gave unequal responses in symmetrically corresponding regions. In some units only two regions could be found, one excitatory and the other inhibitory, lying side by side. In these cases of extreme asymmetry it is possible that there was a second weak flanking area which could not be demonstrated under the present experimental conditions.

An interesting example of a field with only two opposing regions is shown in Fig. 4. The concentrated inhibitory region was confined to an area of about 1° (Fig. 4a). The excitatory area situated to the right of the inhibitory was much larger: a spot of at least 4° was required to evoke a response, and a very strong discharge was seen when the entire 12° excitatory area was illuminated (Fig. 4b). Despite the difference in size between excitatory and inhibitory areas, the effects of stimulating the two together cancelled each other and no response was evoked (Fig. 4c). The semicircular stimulus in Fig. 4b was of special in-

terest because the exact position of the vertical borderline between light and darkness was very critical for a strong response. A slight shift of the boundary to the left, allowing light to infringe on the inhibitory area, completely cancelled the response to illumination. Such a boundary between light and darkness, when properly positioned and oriented, was often an effective type of stimulus.

Cortical receptive fields with central and flanking regions may have either excitatory (Fig. 2) or inhibitory (Figs. 1, 6, 7) centres. So far we have no indication that one is more common than the other.

The axis of a field was defined as a line through its centre, parallel to an optimally oriented elongated stimulus. For each of the field types described, examples were found with axes oriented vertically, horizontally or obliquely. Orientations were determined with respect to the animal's skull. Exact field orientations with respect to the horizontal meridians of the retinas were not known, since relaxation of eye muscles may have caused slight rotation of the eyeballs. Within these limitations the two fields illustrated in Figs. 1-3 were vertically

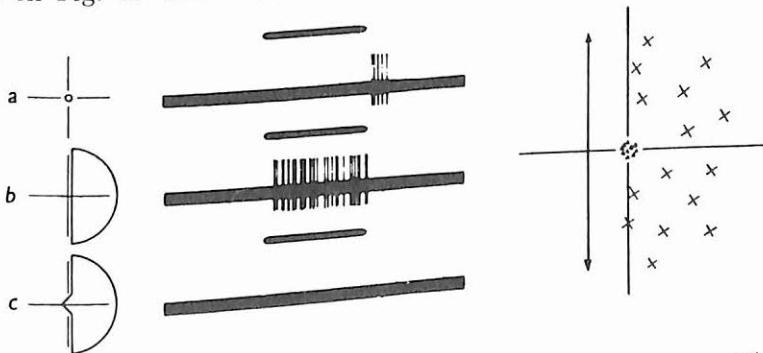


Figure 4

Responses evoked only from contralateral eye. Receptive field just outside nasal border of area centralis. *a*, 1° spot covering the inhibitory region; *b*, right half border of area centralis. *a*, 1° spot covering the inhibitory region; *b*, right half border of area centralis. *c*, light spot covering regions illuminated in *a* and *b*. of a circle 12° in diameter; *c*, light spot covering regions illuminated in *a* and *b*. Background and stimulus intensities and conventions as in Figure 1. Scale, 12° .

arranged: a horizontal field is shown in Figs. 6, 9, and 10, and oblique fields in Figs. 7 and 8.

All units have had their receptive fields entirely within the half-field of vision contralateral to the hemisphere in which they were located. Some receptive fields were located in or near the area centralis, while others were in peripheral retinal regions. All receptive fields were located in the highly reflecting part of the cat's retina containing the tapetum. So far, retinal ganglion cell studies have also been confined to the tapetal region (Kuffler, 1953).

It was sometimes difficult to establish the total size of receptive fields, since the outer borders were often poorly defined. Furthermore, field size may depend on intensity and size of the stimulus spot and on background illumination, as has been shown for the retina by Hartline (1938) and Kuffler (1953). Within these limitations, and under the stimulus conditions specified, fields ranged in total size from about 4° to 10° . Although in the present investigation no systematic studies have been made of changes in receptive fields under different conditions of stimulation, fields obtained in the same unit with the ophthalmoscope and with projection techniques were always found to

be similar in size and structure, despite a difference of several logarithmic units in intensity of illumination. This would suggest that within this photopic range there was little change in size or organization of receptive fields. No units have been studied in states of dark adaptation.

Responses to Movement

Moving a light stimulus in the visual field was generally an effective way of activating units. As was previously found in the freely moving animal (Hubel, 1959), these stimuli were sometimes the only means by which the firing of a unit could be influenced. By moving spots of light across the retina in various directions and at different speeds, patterns of response to movement could be outlined in a qualitative way.

Slit-shaped spots of light were very effective and useful for studies of movement. Here also the orientation of the slit was critical for evoking responses. For example, in the unit of Fig. 3 moving a vertical slit back and forth across the field evoked a definite response at each crossing (Fig. 5*a*), whereas moving a horizontal slit up and down was without effect (Fig. 5*b*). The vertical slit crossed excitatory and inhibitory areas one at a time and

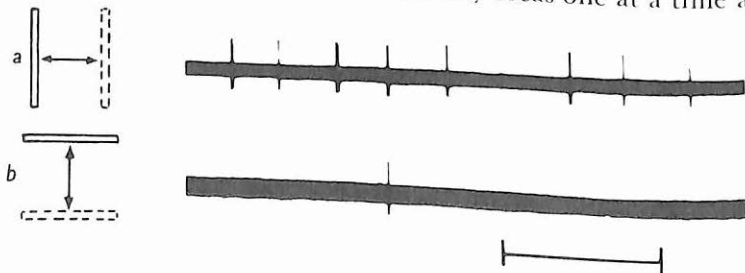


Figure 5

Same unit as in Figures 2 and 3. Receptive field shown in Figure 2. Responses to a slit ($1^\circ \times 8^\circ$) moved transversely back and forth across the receptive field. *a*, slit moved horizontally. *b*, slit moved vertically. Background and stimulus intensities as in Figure 1; time, 1 sec.

each area could exert its effect unopposed, but a horizontal slit at all times covered the antagonistic regions simultaneously, and was therefore ineffective. The response to a vertical slit moved horizontally was about the same for the two directions of movement.

In some units a double response could be observed at each crossing of the receptive field. The receptive field in Fig. 6 had an extensive inhibitory centre flanked by elongated, horizontally oriented, concentrated flanking regions. A horizontal slit moved slowly up or down over the receptive field evoked a discharge as each excitatory region was crossed. A further description of this unit is given in the binocular section of this paper.

Many units showed directional selectivity of a different type in their responses to movement. In these a slit oriented optimally produced responses that were consistently different for the two directions of transverse movement. In the example of Fig. 7, the receptive field consisted of a strong inhibitory area flanked by two excitatory areas, of which the right was weaker than the left. Each region was elongated and obliquely oriented. As usual, a large spot was ineffective

(Fig. 7c). A narrow slit, with its long axis parallel to that of the field, produced a strong response when moved transversely in a direction down and to the left, but only a feeble response when moved up and to the right (Fig. 7d). A tentative interpretation of these findings on the basis of asymmetry within the receptive field will be given in the Discussion.

A number of units responded well to some directions of movement, but not at all to the reverse directions. An example of this is the unit of Fig. 8. Again a slit was moved back and forth transversely in a number of different directions. Only movements up and to the right evoked responses. As with many units, this one could not be activated by stationary stimuli; nevertheless, by using moving stimuli it was possible to get some idea of the receptive field organization—for example, in this unit, the oblique orientation.

Binocular Interaction

Thirty-six units in this study could be driven only from one eye, fifteen from the eye ipsilateral to the hemisphere in which the unit was situated, and twenty-one from the contralateral.

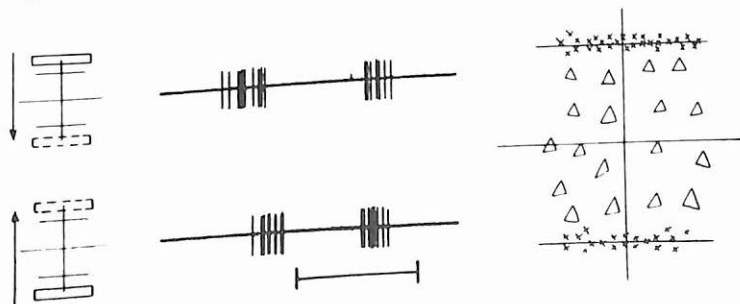


Figure 6

Slow up-and-down movements of horizontal slit ($1^\circ \times 8^\circ$) across receptive field of left eye. Burst of impulses at each crossing of an excitatory region. For details see Figure 9. $\times$, excitatory; Δ , inhibitory. Background illumination -1.9 cd/m^2 ; stimulus intensity 0.69 cd/m^2 ; time, 1 sec.

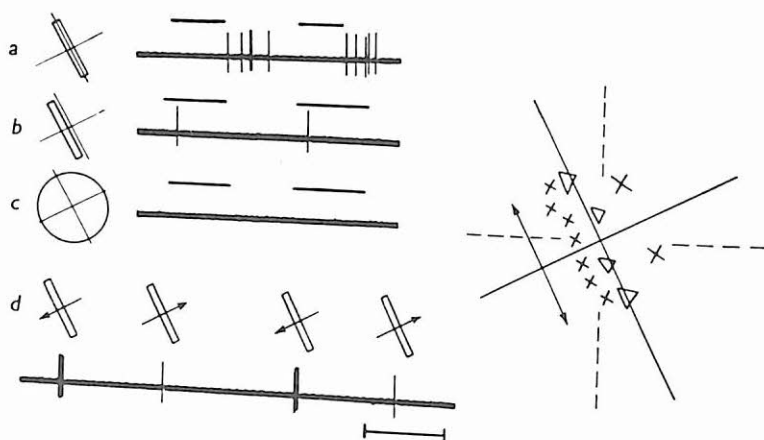


Figure 7
Unit activated from ipsilateral eye only. Receptive field just temporal to area centralis. Field elongated and obliquely oriented. Left excitatory flanking region stronger than right.

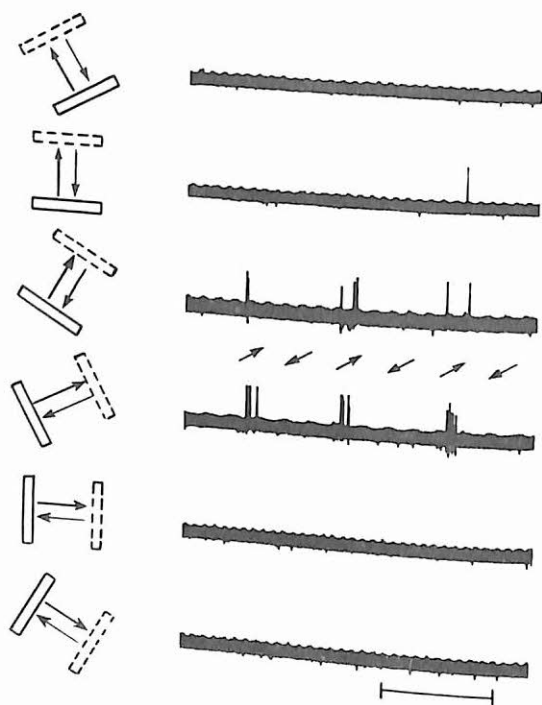


Figure 8
Records from unit activated by ipsilateral eye only; unresponsive to stationary spots, influenced by movement in an area temporal to area centralis. A slit ($0.5^\circ \times 8^\circ$) moved back and forth transversely with different orientations, as shown to the left. For slit orientations evoking responses only one direction was effective—up and to the right. Stimulus and background intensities as in Figure 6; time, 1 sec.

Nine, however, could be driven from the two eyes independently. Some of these cells could be activated just as well from either eye, but often the two eyes were not equally effective, and different degrees of dominance of one eye over the other were seen. In these binocular units the receptive fields were always in roughly homologous parts of the two retinas. For example, a unit with a receptive field in the nasal part of the area centralis of one eye had a corresponding field in the temporal part of the area centralis of the other eye.

Receptive fields were mapped out on a screen in front of the cat. With the eye muscles relaxed with succinylcholine the eyes diverged slightly, so that receptive fields as charted were usually side by side, instead of being superimposed. Whenever the receptive fields of a single unit could be mapped out in the two eyes separately, they were similar in shape, in orientation of their axes, and in arrangements of excitatory and inhibitory regions within the field.

The receptive fields shown in Fig. 9 were obtained from a binocularly ac-

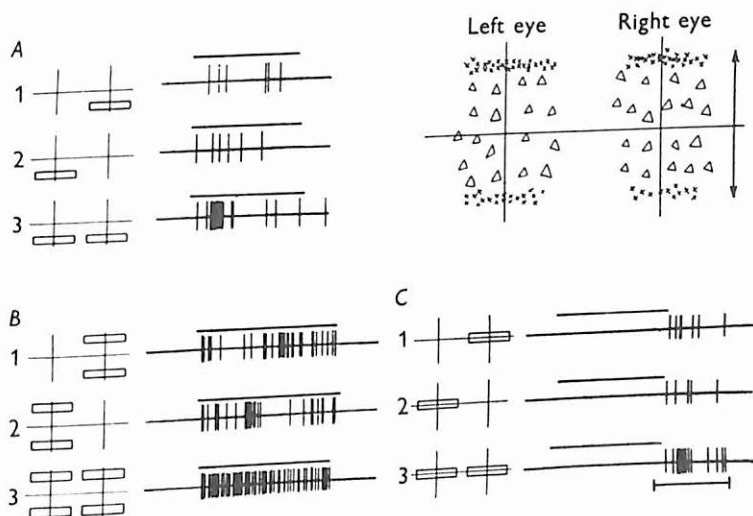


Figure 9

This unit was activated from either eye independently. The illustration shows receptive summation between corresponding parts of the two receptive fields. Receptive field in the contralateral eye was located just above and nasal to area centralis; in the ipsilateral eye, above and temporal. Receptive fields of the two eyes were similar in form and orientation, as shown in upper right of the figure; scale 8° . The pairs of axes in the receptive field diagram are reproduced to the left of each record. Background and stimulus intensities and conventions as in Figure 6. (Same unit as in Figure 6.) A. 1, horizontal slit covering lower flanking region of right eye; 2, same for left eye; 3, pair of slits covering the lower flanking regions of the two eyes. B. 1, pair of horizontal slits covering both flanking regions of the right eye; 2, same for left eye; 3, simultaneous stimulation of all four flanking regions. C. 1, horizontal slit in central region of right eye; 2, same for left eye; 3, simultaneous stimulation of central regions of both eyes. Time, 1 sec.

tivated unit in which each field was composed of an inhibitory centre flanked by narrow horizontal excitatory areas. Responses of the same unit to a horizontal slit moved across the field have already been shown in Fig. 6, for the left eye.

Summation occurred between corresponding regions in the receptive fields of the two eyes (Fig. 9). Thus simultaneous stimulation of two corresponding excitatory areas produced a response which was clearly stronger than when either area was stimulated alone (Fig. 9A). As the excitatory flanks within one receptive field summed, the most powerful response was obtained with a stimulus covering the four excitatory areas in the two eyes (Fig. 9B). Similarly, summation of 'off' responses occurred when inhibitory areas in the two eyes were stimulated together (Fig. 9C).

Antagonism could also be shown between receptive fields of the two eyes (Fig. 10A). Stimulated alone the central area of the left eye gave an 'off'

response, and one flanking area of the right eye gave an 'on' response. When stimulated simultaneously the two regions gave no response. The principles of summation and antagonism could thus be demonstrated between receptive fields of the two eyes, and were not limited to single eyes.

Finally, in this unit it was possible with a moving stimulus to show that opposite-type areas need not always inhibit each other (Fig. 10A), but may under certain circumstances be mutually reinforcing (Fig. 10B). The right eye was covered, and a spot was projected on the screen, over the centre (inhibitory) area of the left eye. Moving the spot as illustrated, away from the centre region of the left eye, produced an 'off' response (Fig. 10B, 1). When the left eye was covered and the right eye uncovered, making the same movement again evoked a response as the flanking excitatory region of the right eye was illuminated (Fig. 10B, 2). The procedure was now repeated with both eyes uncovered, and

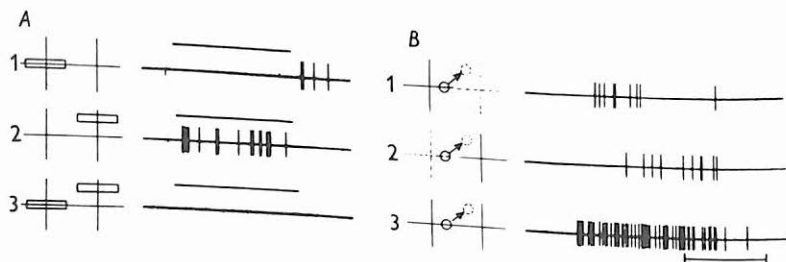


Figure 10

Same unit as in Figure 9. *A*. Antagonism between inhibitory region in the left eye and an excitatory region in the right eye: stationary spots. 1, horizontal slit in centre of left eye; 2, horizontal slit covering upper flanking region of right eye; 3, simultaneous stimulation of the regions of 1 and 2. *B*. Synergism between spot of light. 1, right eye covered, spot moved from inhibitory region in left eye, producing an 'off' response; 2, left eye covered, spot moved into excitatory region in right eye, producing an 'on' response; 3, both eyes uncovered, spot removed from inhibitory region in left eye into excitatory region of right eye, producing a greatly enhanced response. Time, 1 sec.

a greatly increased response was produced. (Fig. 10B, 3). Here the movement was made in such a way that the 'off' response from the left eye apparently added to the 'on' response from the right, producing a response much greater than with either region alone. It is very likely that within a single receptive field opposite-type regions may act in this synergistic way in response to a moving stimulus.

DISCUSSION

In this study most cells in the striate cortex had receptive fields with separate excitatory and inhibitory regions. This general type of organization was first described by Kuffler (1953) for retinal ganglion cells, and has also been found in a preliminary study of neurones in the lateral geniculate body (Hubel & Wiesel, unpublished). Thus at three different levels in the visual system a cell can be inhibited by one type of stimulus and excited by another type, while a stimulus combining the two is always less effective. Most retinal ganglion and geniculate cells give clear responses to a large spot of light covering the entire receptive field. At the cortical level the antagonism between excitatory and inhibitory areas appears to be more pronounced, since the majority of units showed little or no response to stimulation with large spots. Similar findings in the cortex of unanaesthetized, freely moving cats (Hubel, 1959) suggest that this is probably not a result of anaesthesia.

Other workers (Jung, 1953, 1958; Jung & Baumgartner, 1955), using only diffuse light stimulation, were able to drive about half the units in the cat striate cortex, while the remainder could not be activated at all. In recent studies (Hubel, 1960) about

half the units recorded in striate cortex were shown to be afferent fibres from the lateral geniculate nucleus, and these responded to diffuse illumination. The remainder were thought to be cell bodies or their axons; for the most part they responded poorly if at all to diffuse light. The apparent discrepancy between our findings and those of Jung and his co-workers may perhaps be explained by the exclusion of afferent fibres from the present studies. On the other hand it may be that cells responsive to diffuse light flashes are more common in the cortex than our results would imply, but were not detected by our methods of recording and stimulating. However, cortical cells may not be primarily concerned with levels of diffuse illumination. This would be in accord with the finding that in cats some capacity for brightness discrimination persists after bilateral ablation of the striate cortex (Smith, 1937).

The main difference between retinal ganglion cells and cortical cells was to be found in the detailed arrangement of excitatory and inhibitory parts of their receptive fields. If afferent fibres are excluded, no units so far recorded in the cortex have had fields with the concentric configuration so typical of retinal ganglion cells. Moreover, the types of fields found in the cortex have not been seen at lower levels.

Spots of more or less circular (or annular) form are the most effective stimuli for activating retinal ganglion cells, and the diameter of the optimum spot is dependent on the size of the central area of the receptive field (Barlow *et al.*, 1957). At the cortical level a circular spot was often ineffective; for best driving of each unit it was necessary to find a spot with a particular form and orientation. The cortical units described here have had

in common a side-by-side distribution of excitatory and inhibitory areas, usually with marked elongation of one or both types of regions. The form and size of the most effective light stimulus was given by the shape of a particular region. The forms of stimulus used in these studies were usually simple, consisting of slit-shaped spots of light and boundaries between light and darkness. Position and orientation were critical, since imperfectly placed forms failed to cover one type of region completely, thus not taking advantage of summation within that region, and at the same time could invade neighbouring, opposing areas (Fig. 3).

The phenomena of summation and antagonism within receptive fields seem to provide a basis for the specificity of stimuli, in shape, size and orientation. Units activated by slits and boundaries may converge upon units of higher order which require still more complex stimuli for their activation. Most units presented in this paper have had receptive fields with clearly separable excitatory and inhibitory areas. However, a number of units recorded in the striate cortex could not be understood solely in these terms. These units with more complex properties are now under study.

Other types of receptive fields may yet be found in the cortex, since the sampling (45 units) was small, and may well be biased by the micro-electrode techniques. We may, for example, have failed to record from smaller cells, or from units which, lacking a maintained activity, would tend not to be detected. We have therefore emphasized the common features and the variety of receptive fields, but have not attempted to classify them into separate groups.

There is anatomical evidence for convergence of several optic tract fibres on to single geniculate neurons (O'Leary, 1940) and for a more extensive convergence of radiation fibres on to single cortical cells (O'Leary, 1941). Consistent with these anatomical findings, our results show that some single cortical cells can be influenced from relatively large retinal regions. These large areas, the receptive fields, are subdivided into excitatory and inhibitory regions; some dimensions of these may be very small compared with the size of the entire fields. This is illustrated by the fields shown in Figs. 1, 2 and 7, in which the central regions were long but very narrow; and by that of Fig. 9, in which both flanks were narrow. It is also shown by the field of Fig. 4, which had a total size of about 12° but whose inhibitory region was only about 1° in diameter. Thus a unit may be influenced from a relatively wide retinal region and still convey precise information about a stimulus within that region.

Movement of a stimulus across the retina was found to be a very potent way of activating a cell, often more so than a stationary spot. Transverse movement of a slit usually produced responses only when the slit was oriented in a certain direction. This was sometimes explained by the arrangement within the receptive fields as mapped out with stationary stimuli (Fig. 5).

In many units (Fig. 7) the responses to movement in opposite directions were strikingly different. Occasionally when the optimum direction of movement was established, there was no response to movement in the opposite direction (Fig. 8). Similar effects have been observed with horizontally moving spots in the unanaesthetized

animal (Hubel, 1959). It was not always possible to find a simple explanation for this, but at times the asymmetry of strength of flanking areas was consistent with the directional specificity of responses to movement. Thus in the unit of Fig. 7 best movement responses were found by moving a slit from the inhibitory to the stronger of the two excitatory regions. Here it is possible to interpret movement responses in terms of synergism between excitatory and inhibitory areas. This is further demonstrated in Fig. 10 *B*, where areas antagonistic when tested with stationary spots (Fig. 10 *A*) could be shown to be synergistic with moving stimuli, and a strong response was evoked when a spot moved from an 'off' to an 'on' area.

Inhibition of unitary responses by stimulation of regions adjacent to the excitatory area has been described for the eccentric cell in the *Limulus* eye (Hartline, 1949) and for ganglion cells both in the frog retina (Barlow, 1953) and in the cat retina (Kuffler, 1953). Analogous phenomena have been noted for tones in the auditory system (dorsal cochlear nucleus, Galambos, 1944) and for touch and pressure in the somatosensory system (Mountcastle, 1957). In each system it has been proposed that these mechanisms are concerned with enhancing contrast and increasing sensory discrimination. Our findings in the striate cortex would suggest two further possible functions. First, the particular arrangements within receptive fields of excitatory and inhibitory regions seem to determine the form, size and orientation of the most effective stimuli, and secondly, these arrangements may play a part in perception of movement.

It is clear from stimulation of sepa-

rate eyes with spots of light that some cortical units are activated from one eye only, either the ipsilateral or the contralateral, while others can be driven by the two eyes. In view of the small number of cells studied, no conclusion can be drawn as to the relative proportions of these units (ipsilaterally, contralaterally and bilaterally driven), but it appears that all three types are well represented.

Studies of binocularly activated units showed that the receptive fields mapped out separately in the two eyes were alike. The excitatory and inhibitory areas were located in homologous parts of the retinas, were similarly shaped and oriented, and responded optimally to the same direction of movement. When corresponding parts of the two receptive fields were stimulated summation occurred (Fig. 9). Assuming that the receptive fields as projected into the animal's visual field are exactly superimposed when an animal fixes on an object, any binocularly activated unit which can be affected by the object through one eye alone should be much more strongly influenced when both eyes are used. The two retinal images of objects behind or in front of the point fixed will not fall on corresponding parts of the fields, and their effects should therefore not necessarily sum. They may instead antagonize each other or not interact at all.

It is possible that when an object in the visual field exerts, through the two eyes, a strong influence on binocularly activated units, those influences may lead in some way to an increased awareness of the object. If that is so, then objects which are the same distance from the animal as the object fixed should stand out in relief. On the other hand such units may be

related to mechanisms of binocular fixation, perhaps projecting to mid-brain nuclei concerned with the regulation of convergence.

SUMMARY

1. Recordings were made from single cells in the striate cortex of lightly anaesthetized cats. The retinas were stimulated separately or simultaneously with light spots of various sizes and shapes.

2. In the light-adapted state cortical cells were active in the absence of additional light stimulation. Increasing the depth of anaesthesia tended to suppress this maintained activity.

3. Restricted retinal areas which on illumination influenced the firing of single cortical units were called receptive fields. These fields were usually subdivided into mutually antagonistic excitatory and inhibitory regions.

4. A light stimulus (approximately 1 sec duration) covering the whole receptive field, or diffuse illumination of the whole retina, was relatively ineffective in driving most units, owing to mutual antagonism between excitatory and inhibitory regions.

5. Excitatory and inhibitory regions, as mapped by stationary stimuli, were arranged within a receptive field in a side-by-side fashion with a central area of one type flanked by antagonistic areas. The centres of receptive fields could be either excitatory or inhibitory. The flanks were often asymmetrical, in that a given stationary stimulus gave unequal responses in corresponding portions of the flanking areas. In a few fields only two regions could be demonstrated, located side-by-side. Receptive fields could be oriented in a vertical, horizontal or oblique manner.

6. Effective driving of a unit re-

quired a stimulus specific in form, size, position and orientation, based on the arrangement of excitatory and inhibitory regions within receptive fields.

7. A spot of light gave greater responses for some directions of movement than for others. Responses were often stronger for one direction of movement than for the opposite; in some units these asymmetries could be interpreted in terms of receptive field arrangements.

8. Of the forty-five units studied, thirty-six were driven from only one eye, fifteen from the ipsilateral eye and twenty-one from the contralateral; the remaining nine could be driven from the two eyes independently. In some binocular units the two eyes were equally effective; in others various degrees of dominance of one eye over the other were seen.

9. Binocularly activated units were driven from roughly homologous regions in the two retinas. For each unit the fields mapped for the two eyes were similar in size, form and orientation, and when stimulated with moving spots, showed similar directional preferences.

10. In a binocular unit excitatory and inhibitory regions of the two receptive fields interacted, and summation and mutual antagonism could be shown just as within a single receptive field.

We wish to thank Dr. S. W. Kuffler for his helpful advice and criticism, and Mr. R. B. Bosler and Mr. P. E. Lockwood for their technical assistance. This work was supported in part by U.S. Public Health Service grants B-22 and B-1931, and in part by U.S. Air Force contract AF 49 (638)-499 (Air Force Office of Scientific Research, Air Research and Development Command).

REFERENCES

- BARLOW, H. B. (1953). Summation and inhibition in the frog's retina. *J. Physiol.* **119**, 69-88.
- BARLOW, H. B., FITZHUGH, R., & KUFFLER, S. W. (1957). Change of organization in the receptive fields of the cat's retina during dark adaptation. *J. Physiol.* **137**, 338-354.
- DAVIES, P. W. (1956). Chamber for micro-electrode studies in the cerebral cortex. *Science*, **124**, 179-180.
- FRANK, K. & FUORTES, M. G. F. (1955). Potentials recorded from the spinal cord with micro-electrodes. *J. Physiol.* **130**, 625-654.
- GALAMBOS, R. (1944). Inhibition of activity in single auditory nerve fibers by acoustic stimulation. *J. Neurophysiol.* **7**, 287-304.
- HARTLINE, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.* **121**, 400-415.
- HARTLINE, H. K. (1949). Inhibition of activity of visual receptors by illuminating nearby retinal areas in the *Limulus* eye. *Fed. Proc.* **8**, 69.
- HUBEL, D. H. (1957). Tungsten micro-electrode for recording from single units. *Science*, **125**, 549-550.
- HUBEL, D. H. (1959). Single unit activity in striate cortex of unrestrained cats. *J. Physiol.* **147**, 226-238.
- HUBEL, D. H. (1960). Single unit activity in lateral geniculate body and optic tract of unrestrained cats. *J. Physiol.* (In the Press.)
- JUNG, R. (1953). Neuronal discharge. *Electroenceph. clin. Neurophysiol. Suppl.* **4**, 57-71.
- JUNG, R. (1958). Excitation, inhibition, and coordination of cortical neurones. *Exp. Cell Res. Suppl.* **5**, 262-271.
- JUNG, R. & BAUMGARTNER, G. (1955). Hemmungsmechanismen und bremsende Stabilisierung an einzelnen Neuronen des optischen Cortex. *Pflüg. Arch. ges. Physiol.* **261**, 434-456.
- KUFFLER, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.* **16**, 37-68.
- MOUNTCASTLE, V. B. (1957). Modality and topographic properties of single neurons of cats' somatic sensory cortex. *J. Neurophysiol.* **20**, 408-434.
- O'LEARY, J. L. (1940). A structural analysis of the lateral geniculate nucleus of the cat. *J. comp. Neurol.* **73**, 405-430.
- O'LEARY, J. L. (1941). Structure of the area striata of the cat. *J. comp. Neurol.* **75**, 131-164.
- SMITH, K. U. (1937). Visual discrimination in the cat: V. The postoperative effects of removal of the striate cortex upon intensity discrimination. *J. genet. Psychol.* **51**, 329-369.
- TALBOT, S. A. & MARSHALL, W. H. (1941). Physiological studies on neural mechanisms of visual localization and discrimination. *Amer. J. Ophthalm.* **24**, 1255-1263.
- TALBOT, S. A. & KUFFLER, S. W. (1952). A multibeam ophthalmoscope for the study of retinal physiology. *J. opt. Soc. Amer.* **42**, 931-936.

INTRODUCTION

When you move your eyes along the horizon all the objects in the field of view are sweeping over your retina. When you fixate your eyes and the visual scene in front of you is moving, all objects in the field of view are also sweeping over your retina. In the first instance, you do not perceive the world as moving, while in the second you do. Why? What's the difference? The difference is that when a person directs his eyes to move he puts into his brain what is called a motor-corollary discharge. It is a piece of information that tells the brain "Get ready: Your eyes are about to move and objects will appear to be moving across your retina. Don't pay any attention to them, and consider them to be held constant." It is this mysterious signal that is responsible for stability of much of our visual world. It must physically exist somewhere, but to date has eluded most experimenters. Consider the problem. It is fascinating.

RELATIONS BETWEEN THE CENTRAL NERVOUS SYSTEM AND THE PERIPHERAL ORGANS

E. von Holst

The relation of the Central Nervous System (CNS) to the peripheral senses and muscular movement is an old and much discussed problem. Here we are at the heart of the physiology of behaviour, and in comparison to that which is not known, our present knowledge is very meagre and vague! Under these circumstances, our knowledge and conceptions are dependent upon the method which happens to be popular at the moment. In this field, the method which has played the greatest role consists of, first, artificially inactivating the CNS and then, through peripheral stimulation, evoking a particular response. On this basis, the CNS is often held to be only a reflex-mechanism, yet we know today that this view is one-sided. In order to be in co-ordinated activity, the CNS often needs a minimum of stimulation or loading by afferent impulses; the conception of chain-reflex-co-ordination has been recognised almost everywhere as being incorrect. Isolated, that is de-afferented, parts of

the nervous system show continued electrical activity. One can therefore say that, as a rule, de-afferented ganglion cells, under otherwise normal conditions, possess "automaticity."

These facts allow us to regard the function of the peripheral senses from a new viewpoint. The classical reflex-concept assumes that the peripheral stimulus initiates the central nervous activity. Since we now know that this supposed cause is often unnecessary, it is possible to start from the CNS. We can ask the question, what effect is produced on the sensory-receptors by the motor impulses which initiate a muscular movement? Thus, we look from the opposite direction, not from the outside inward, but from the centre to periphery. You will quickly see that in this manner we shall come upon new problems and experimentally verifiable hypotheses.

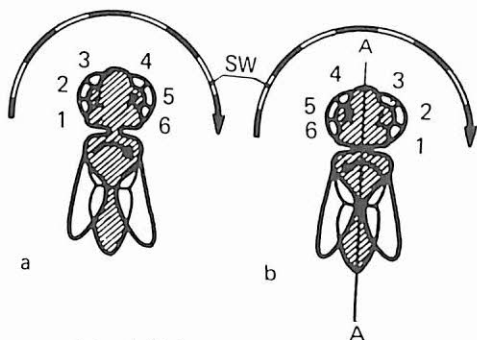
In order to make myself clear, I should like first to explain a few terms. The whole of the impulses which are produced by whatever stimuli in whatever receptors I shall term *afference*, and in contradistinction to this I shall call the whole of the motor impulses *efference*. Efference can only be present when ganglion cells are

We are grateful to Mrs. Eveline von Holst and the publisher for their kind permission to include this selection. Reprinted from the British Journal of Animal Behavior, 1954, pp. 89-94.

active; afference, on the contrary, can have two quite different sources: first, stimuli produced by muscular activity, which I shall call *re*-afference; second, stimuli produced by external factors, which I shall call *ex*-afference. *Re*-afference is the necessary afferent reflexion caused by every motor impulse; *ex*-afference is independent of motor impulses.

Here are some examples: when I turn my eyes, the image present on the retina moves over the retina. The stimuli so produced in the optic nerve constitute a *re*-afference, for this is the necessary result of my eye movement. If I shake my head, a *re*-afference necessarily is produced by the labyrinth. If, on the other hand, I stand on a railway platform looking straight at a train when it starts to move, the moving image on the retina of my unmoving eye produces an *ex*-afference; likewise, when I lie in a tossing ship, the impulses of my labyrinth will constitute an *ex*-afference. If I shake the branch of a tree, various receptors of my skin and joints produce a *re*-afference, but if I place my hand on a branch shaken by the wind, the stimuli of the same receptors produce an *ex*-afference. We can see that this distinction has nothing to do with the difference between the so called proprio- and extero-receptors. The *same* receptor can serve both the *re*- and the *ex*-afference. The CNS must, however, possess the ability to distinguish one from the other. This distinction is indispensable for every organism, since it must correctly perceive its environment at rest and in movement, and stimuli resulting from its own movements must not be interpreted as movements of the environment. I want to describe experiments which show how the CNS distinguishes between *ex*-afference and *re*-afference.

When one rotates a striped cylinder around a quietly sitting insect, for instance the fly *Eristalis*, the animal turns itself in the same sense (Fig. 1a).



Insect (*Eristalis*) in striped cylinder (SW), L, R=left, right eyes; a=head in normal, b=in turned position.

Figure 1

Insect (*Eristalis*) in striped cylinder (SW), L, R = left, right eyes; a = head in normal, b = turned position.

This is a well-known optomotor-“reflex.” As soon as the animal moves itself, for instance, “spontaneously” (or stimulated by a smell), one observes that it turns itself unhindered by the stripes of the stationary cylinder. We must ask ourselves why the animal at every turn is not turned back by his optomotor “reflex,” since the movement of the image on the retina is the same as in the first case, when the cylinder moved and the animal was stationary. A possible answer according to the reflex-theory is that in locomotion the optomotor-“reflex” is inhibited or “blocked.” But we shall see that this answer is incorrect. It is possible, as has been shown by my colleague Mittelstaedt, to turn the head of the insect through 180° about the long axis (Fig. 1b A-A); then the head is fixed to the thorax, so that the two eyes are effectively interchanged and the order

of the visual elements is reversed. The unmoving animal now responds, when the cylinder turns to the right, by turning itself to the left, as is to be expected from the reversed position of the eyes. If it is indeed the case that in spontaneous (or otherwise caused) locomotion the optomotor-"reflex" is "blocked," the animal should move unhindered in the stationary cylinder. But the opposite is the case; once the insect begins to move, it spins rapidly to right or left in small circles until it is exhausted. We have observed the same behaviour with fishes, whose eyes have been turned 180° about the optic axis. But we have found this behaviour only in patterned optical surroundings; in optically homogeneous surroundings the animal moves normally. This indicates that the optomotor-"reflex" is not "blocked" in locomotion, but on the contrary, the associated re-afference plays an important role. Exactly what that role is will be made clearer by the next example.

If a vertebrate is turned over on its side by external forces, the well-known righting "reflexes" are initiated by the ex-afference of the labyrinth. But, just as in my first example, every animal is able to take up any position without righting reflexes being produced by the re-afference of the labyrinth. Again, it has been believed that the reflexes were "blocked" during position changing; and, again, we can show that this is not the case.

The righting reflexes, as is well-known, are released by the statoliths in the labyrinths, which, when the head is tilted, produce a shearing force on the underlying sensory organ, as we have found in fishes. One can increase this mechanical force which the statoliths exert on the sense organs, through the addition of a con-

stant centrifugal force. We have built for this purpose a small revolving laboratory, capable of more than doubling the gravitational force. In this manner the statolith is made heavier, and the corresponding shearing stimuli produced by every tilting of the head are quantitatively increased. If one records the tilting of free swimming fish under these conditions, one finds that the degree of tilting becomes proportionally less, the heavier the statoliths are made. (For the method of measurement see v. Holst u. Mittelstaedt, 1950). If the statoliths are removed, then the behaviour of the fish is the same under normal and centrifugal conditions. We see, therefore, that the re-afference of the labyrinth is not "blocked", but has a quantitative effect upon the *degree* of tilting, and, indeed, the greater the re-afference, the smaller the degree of the movement. One can say that the CNS "measures" the degree of movement by the magnitude of the re-afference thereby released.

Thus we have learned two facts: if the form of the re-afference is reversed, as in the first example, then the initiated movement is increased progressively. Secondly, if the re-afference keeps its normal form but is increased, as in the second example, the initiated movement is correspondingly decreased. These facts allow us to formulate a hypothesis about the mechanism here involved. We shall propose that the efference leaves an "image" of itself somewhere in the CNS, to which the re-afference of this movement compares as the negative of a photograph compares to its print; so that, when superimposed, the image disappears. Figure 2 illustrates this in a number of subsequent steps. A motor impulse, a "command" C (Fig. 2a), from a higher centre HC

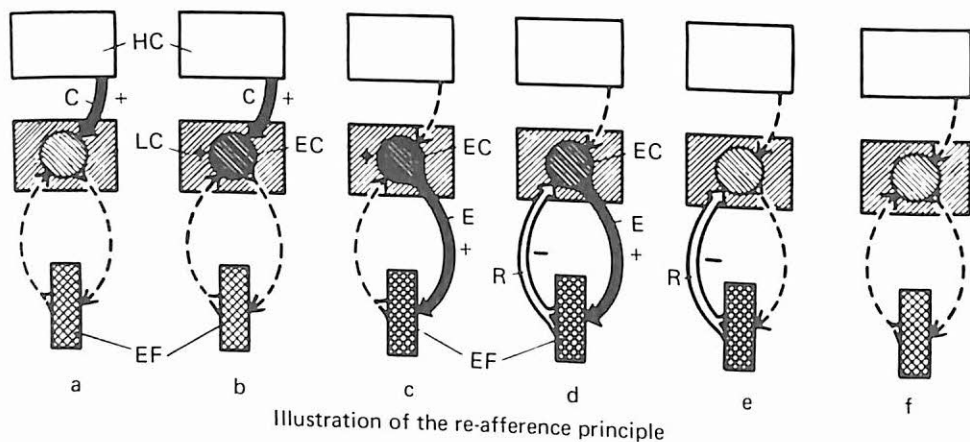


Illustration of the re-reference principle

Illustration of the re-reference principle; see explanation in text.

Figure 2

causes a specific activation in a lower centre LC (Fig. 2b), which is the stimulus-situation giving rise to a specific efference E (Fig. 2c) to the effector EF (i.e. a muscle, a joint, or the whole organism). This central stimulus-situation, the "image" of the efference, may be called "efference copy," EC. The effector, activated by the efference, produces a re-efference R, which returns to the lower centre, nullifying the efference copy by superposition (Fig. 2d-f). Because of the complementary action of these two components we can arbitrarily designate the whole efferent part of this process as plus (+, dark coloured) and the afferent part as minus (-, white coloured). When the efference copy and the re-efference exactly compensate one another, nothing further happens. When, however, the efference is too small or lacking, then a + difference will remain or when the re-efference is too great, a - difference will remain. This difference will have definite effects, according to the particular organisation of the system. The difference can either influence the movement itself, or for instance,

ascend to a higher centre and produce a perception.

Let us first consider the simple situation of Fig. 2. The initiated movement will continue, until the re-efference exactly nullifies the efference copy. Then we must predict the following: first, if through external influence the re-efference is increased, then the initiated movement will end prematurely. We have already seen that this is the case in the fish labyrinth experiment with the centrifuge. Secondly (Fig. 3a), if the re-efference is inverted, that is changed from - to +, there will be no nullification, but summation (Fig. 3b) and the movement will progressively increase, as we have already seen in the experiment with the inverted eyes.* Thirdly (Fig. 4), in the case where the re-efference is lacking (for instance, due to the destruction of the afferent pathways) the initiated movement will not be increased, as in the second case, but will continue until something else limits it. This behaviour occurs widely and can be seen particularly well in

* This is the so called "positive feed-back."

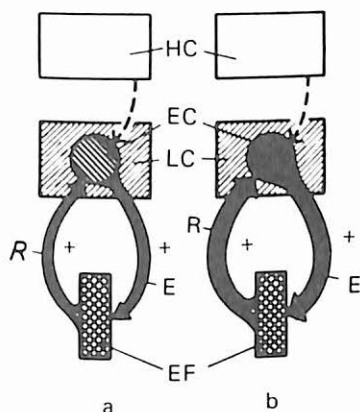


Illustration of the experiment with the eyes in turned position

Figure 3

Illustration of the experiment with the eyes in turned position (Figure 1); see explanation in text.

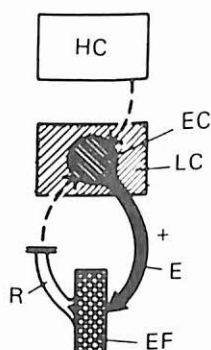


Illustration of the experiment with interrupted

Figure 4

Illustration of the experiment with interrupted efference; see explanation in text.

fish without labyrinths in optically homogeneous surroundings. Every turning or tilt leads to circling or somersaulting. Also, in the human disease *Tabes dorsalis*, where the dorsal roots are destroyed, the well known exaggerated, ataxic movements of the limbs indicate that the same mechanism is involved. Therefore, contrary to the chain-reflex theory, the stimulus, originating with every movement, that is the re-affectance, produces

not an augmenting, excitatory, but a *limiting*, effect on the movement. Only those forms of locomotion, such as the swimming of fish, which do not require a constant adjustment to the surrounding medium, proceed just as before after de-affectation. These movements are automatically co-ordinated in the CNS and therefore require no limiting re-affectance (v. Holst, Lissmann).

With this simple scheme we are able to understand a number of previously unexplained types of behaviour. The most hypothetical part of this theory is the postulated efference copy: this "image" in the CNS, produced by the "command" and matched by the re-affectance. I am going to present direct proof of the existence of this phenomenon. For this purpose I choose two human examples, in which the difference between the efference copy and re-affectance is transmitted to a higher centre and produces a perception. My first example is concerned with the already mentioned human eye movement.

A re-affectance from the actively moving eye can have two sources: firstly, movement of the image across the retina and secondly, impulses from the sensory cells of the eye muscles. The former results in a conscious perception; the latter is of no importance for the following consideration. Consider my eye mechanically fixed and the muscle receptors narcotised (Fig. 5a). When I want to turn my eye to the right, an efference *E* and, according to the theory, an efference-copy *EC* is produced, but the immovable eye does not produce any re-affectance. The efference-copy will not be nullified, but transmitted to higher centres and could produce a perception. It is possible to predict the exact form of this perception (v.

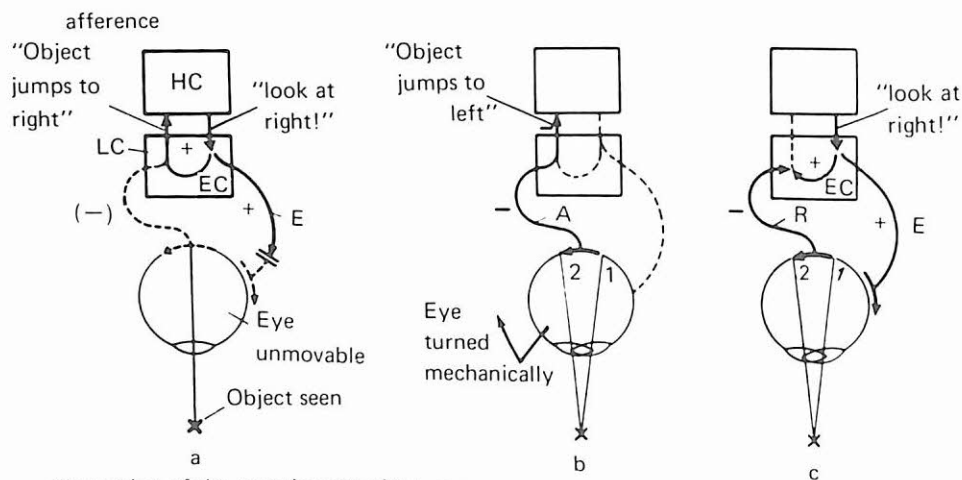


Illustration of the experiments with human eye

Figure 5

Illustration of the experiments with human eye; explanation in text (for the letters compare the text of Figure 2).

Holst und Mittelstaedt, 1950). The perception, if I want to turn my eye to the right, must be that "the surroundings have jumped to the right." This is indeed the case! It has been known for many years from people with paralysed eye muscles and it has been established exactly from the experiments of Kornmüller on himself that every intended but unfulfilled eye movement results in the perception of a quantitative movement of the surroundings in the same direction. Since here *nothing* happens on the afferent pathways, this false perception *can* only result from the activity, originated by the intention of the eye movement, being returned to higher centres. This is another way of saying that the unmatched efference-copy causes the perception.

Now, we make a simple experiment and turn the paralysed eye mechanically to the right (Fig. 5b). In this case both the motor intention and also the efference-copy are lacking, but the image moves across the retina

and afference A is transmitted, unmatched by an efference-copy, to higher centres and produces, as is known, the perception that "the surroundings move to the left." This is also a false perception. If now we combine the first case with the second, that is, if my eye is moved mechanically at the same time I intend this movement—which is the same as *voluntarily* moving a *normal* eye—then in fact these two complementary effects just mentioned are produced: firstly, the perception of the returning "command" causing a jump of the surroundings to the right and, secondly, an image-motion on the retina producing a jump of the surroundings in the opposite direction. These two phenomena, the efference-copy and the re-efference, now compensate each other (Fig. 5c); and as a result *no moving* of the surroundings is perceived. The surroundings appear stationary during this normal eye movement, and *this* perception is *physically correct*. As we have already

seen, the correct perception results from two opposite and false perceptions which cancel each other. Thus, we understand a phenomenon with which Psychology has been concerned for many years, that is, the perception of the surroundings as nearly stationary during eye movements ("Raumkonstanz").

Now we come to the second example, visual accommodation. The eye is focussed for distant vision when at rest, since the elastic lens is flattened by its zonal fibres. For near-accommodation a circular muscle, working against these fibres, allows the lens to round up. We should also like to apply our theory to this system. If the accommodation apparatus is narcotised (for instance by atropine), that is, the eye is permanently accommodated for distant vision, then an intention for near-accommodation will start a motor-impulse, which cannot be nullified by any re-affectance and, therefore, must return to a higher centre, where it can produce a perception. This is indeed the case. All objects in the visual field become small, and this false perception is called "micropsia." The same phenomenon must exist with a normal eye, if we imprint an after-image of a distant cross on the retina and then look upon a near surface. Since the after-image remains the same size and sharpness on the retina, it must appear very small on the near surface, because again only the "command" for accommodation returns to the centre of perception. This is also the case, as one can easily convince oneself. These false perceptions appear, although the *peripheral stimulus-situation* is unaltered. If, on the contrary, the accommodation of a normal eye is unaltered, that is, if we look first at a small and then at a large cross at the

same distance, then naturally the changed affectance will be transmitted to the centre of perception and we see the second cross to be larger. Now we combine this last case with the first, that is, we observe with a *normal* eye a cross, moving from a distant point nearer to the eye. This initiates the accommodation-impulse, which returning, tells us "the cross is becoming smaller"; but at the same time the enlargement of the retinal image states, "the cross is becoming larger." The two cancel one another out, with the result that we perceive the cross to be of *constant* size. Again, the correct perception is the result of two opposite false perceptions; and, further, we come to an understanding of a phenomenon, long discussed in Psychology, the "Grossen Konstanz der Seh-Dinge" (Hering), which means that we see the objects to be nearly the same size irrespective of their distance from us.

I could present still further examples from man and from lower and higher animals which would show what role the re-affectance plays in general in behaviour. It serves either to limit the magnitude of movement or to insure the constancy of the perceived surroundings during movement, and so makes possible the distinction between real and apparent motion of objects. The first step in both of these functional mechanisms is the comparison of the re-affectance with the efference-copy.

In conclusion, permit me a few general considerations. I have attempted to show through the example of this central nervous mechanism, that it is possible in the field of the Physiology of Behaviour to avoid formulating "theories," which are only generalised descriptions of observations; rather should we follow the example of the

exact sciences, namely, that a theory must exactly predict what will happen under defined conditions, so that one can by experiment verify or disprove it. Thus one avoids the error of false generalisation, which often occurs in central nervous physiology. For this reason I would like to emphasize that the principle of re-reference is only one of many central nervous mechanisms.

There exists a large number of other mechanisms with other modes of function, and of these we know as yet very little. We recognise fragments of some of them and call them "reflexes"; but this term denotes fragments of very different mechanisms. I believe the whole Central Nervous system is a "hierarchical system" of such different functional parts, a concept which you find also in Tinbergen's book "The Study of Instinct."

One final point. I have spoken of neither electrical spikes, nor nerve pathways, nor anatomical centres, in which particular functions might be localised. In the realm of behavioural analysis these things are indeed of secondary interest. The functional

schemata, constructed in order to illustrate definite causal relationships, are quite abstract, although the consequences they predict are concrete and experimentally verifiable. The physiologist who fully understands such a causal system is still unable to deduce where the cell elements which perform this function are located, or how they operate. Such questions are dealt with at another level of investigation, where the electrophysiologist works and develops his own terminology. It is useful and justifiable for every level of investigation to have its own language, but we must expect that, with a greater advancement of our knowledge, it will be easy to translate one such language into another. Until such a time, each field must develop along its own lines, unhindered by the many possibilities for misinterpretation.

Reference

- v. Holst u. Mittelstaedt (1950). *Naturwissenschaften*, 464-476.

INTRODUCTION

The discovery that there are specific sites in the brain that produce behavioral effects of punishment and reward came as a dramatic surprise to many psychologists in the fifties. By implanting electrodes at various points in the brain, Olds and Milner were able to show that when current was passed through the electrodes, the rats would act as though the stimulation were rewarding in some instances and punishing in others. A thousand questions were immediately precipitated by this work. Were pleasurable events experienced by normal humans nothing other than the rapid discharge of a particular brain site at a given moment? These and a host of other questions are raised by these fascinating studies. What do they suggest to you?

SELF-STIMULATION OF THE BRAIN

James Olds

This article reviews experiments which have led to the discovery and analysis of localized systems in the brain where electric stimulation has positive and negative motivational effects (1). Basically, the experimental animal in these studies is rewarded or punished by a brain shock. The site of electric stimulation determines the motivational effect.

Reprinted with permission from Science, 127, February 14, 1958, pp. 316-24.

The studies are important primarily as a beginning step toward filling the large gap which has existed between neurophysiological techniques and an understanding of complex psychological processes. Among other things, they carry the enterprise of brain mapping into the realm of clearly defined motivational functions; this by itself correlates an orderly array of integrative psychological mechanisms with an orderly array of anatomical

exact sciences, namely, that a theory must exactly predict what will happen under defined conditions, so that one can by experiment verify or disprove it. Thus one avoids the error of false generalisation, which often occurs in central nervous physiology. For this reason I would like to emphasize that the principle of re-afference is only *one* of *many* central nervous mechanisms. There exists a large number of other mechanisms with other modes of function, and of these we know as yet very little. We recognise fragments of some of them and call them "reflexes"; but this term denotes fragments of very different mechanisms. I believe the whole Central Nervous system is a "hierarchical system" of such different functional parts, a concept which you find also in Tinbergen's book "The Study of Instinct."

One final point. I have spoken of neither electrical spikes, nor nerve pathways, nor anatomical centres, in which particular functions might be localised. In the realm of behavioural analysis these things are indeed of secondary interest. The functional

schemata, constructed in order to illustrate definite causal relationships, are quite abstract, although the consequences they predict are concrete and experimentally verifiable. The physiologist who fully understands such a causal system is still unable to deduce where the cell elements which perform this function are located, or how they operate. Such questions are dealt with at another level of investigation, where the electrophysiologist works and develops his own terminology. It is useful and justifiable for every level of investigation to have its own language, but we must expect that, with a greater advancement of our knowledge, it will be easy to translate one such language into another. Until such a time, each field must develop along its own lines, unhindered by the many possibilities for misinterpretation.

Reference

- v. Holst u. Mittelstaedt (1950). *Naturwissenschaften*, 464-476.

INTRODUCTION

The discovery that there are specific sites in the brain that produce behavioral effects of punishment and reward came as a dramatic surprise to many psychologists in the fifties. By implanting electrodes at various points in the brain, Olds and Milner were able to show that when current was passed through the electrodes, the rats would act as though the stimulation were rewarding in some instances and punishing in others. A thousand questions were immediately precipitated by this work. Were pleasurable events experienced by normal humans nothing other than the rapid discharge of a particular brain site at a given moment? These and a host of other questions are raised by these fascinating studies. What do they suggest to you?

SELF-STIMULATION OF THE BRAIN

James Olds

This article reviews experiments which have led to the discovery and analysis of localized systems in the brain where electric stimulation has positive and negative motivational effects (1). Basically, the experimental animal in these studies is rewarded or punished by a brain shock. The site of electric stimulation determines the motivational effect.

Reprinted with permission from Science, 127, February 14, 1958, pp. 316-24.

The studies are important primarily as a beginning step toward filling the large gap which has existed between neurophysiological techniques and an understanding of complex psychological processes. Among other things, they carry the enterprise of brain mapping into the realm of clearly defined motivational functions; this by itself correlates an orderly array of integrative psychological mechanisms with an orderly array of anatomical

points in the brain. Furthermore, these studies perform a unification long considered technically impossible between electrophysiological, independent variables and standard, behavioral, dependent variables to produce smooth interaction curves relating the two.

For psychologists, these experiments help to clarify the basic notions of reward and punishment. Reward and punishment, it is agreed, determine which behaviors will predominate in an organism's repertory and which will be erased from it. Rewarded responses are repeated more frequently than would be expected by chance; punished ones are repeated with less frequency. This is obvious.

Less self-evident is the thesis of the classical theory of reward, according to which reward is interpreted as being the falling phase of the same massive stimulation which at high levels constitutes punishment. This thesis is greatly weakened by the work outlined in this article; however, it has held sway for such a long time in psychology and conditions so many basic attitudes that it will certainly form a foundation stone for the new theories which replace it.

Drive and punishment are synonymous, according to this theory, and a reward is held to be fundamentally nothing more than the reduction of a drive. Physiological conditions which are inimical to survival, such as food deficiencies or tissue damage, cause massive receptor discharge into the central nervous system. This discharge is held to be the drive, and it is held to be reflected in behavioral activation. The latter is a nonselective function of the massive drive stimulation, energizing adaptive and maladaptive responses equally. The drive stimulation, however, also has a selective

function by which it combines with other cue stimuli to select those responses which have repaired this particular physiological deficit in the past.

The response actually selected by the combination of drive and cue stimuli is determined entirely by structural cue-response connections whose strength has been determined by prior learning. More specifically, a group of cues actually selects the response which was previously followed by drive reduction in their presence. The drive reduction, on previous occasions, caused a rewarding or positive reinforcing effect which somehow increased the causal connection between these stimuli and this particular response.

The hedonistic view that behavior is pulled forward by pleasure as well as pushed forward by pain is rejected in this classical theory for the more parsimonious notion that pain supplies the push and that learning based on pain reduction supplies the direction.

The work reported in this article clearly shows one implication of the drive-reduction theory to be incorrect, for massive inputs to certain parts of the central nervous system are shown to have rewarding effects. Further, by showing that there are anatomically separate mechanisms for reward and punishment in the brain, it points directly to a physiological basis for the motivational dualism suggested in the hedonistic theory.

In fact, it appears that the area producing rewarding effects, upon electric stimulation, is far larger than the area producing punishment. In one early experiment, 76 electrodes were implanted in the brains of rats in an attempt to get a random sampling of midbrain and forebrain points. It was found that stimulation at 47 of these

points had motivational effects. Stimulation at 36 of these motivational points produced approach behavior—that is, the rats stimulated themselves repeatedly by means of the technique described below; at only 11 points did stimulation produce avoidance behavior (2).

Basic Studies

METHOD

The method of self-stimulation (3) is modeled in part after the chronic implantation technique of W. R. Hess (4) and in part after the box technique of B. F. Skinner (5). The former developed a technique for implanting electrodes permanently in the brain; the technique allows stimulation in the freely behaving animal. The latter worked out a way to measure positive reinforcement—that is, reward—by arranging a situation in which the experimental animal could deliver the reward to itself by a very simple manipulation, and then counting the frequency of the manipulations. Self-stimulation combines these techniques by allowing animals to deliver shocks to specific points in their brains through chronically implanted electrodes.

Figure 1 shows the method used. When the rat stepped on a pedal, a shock was delivered to its brain. The rat never received any other reward for pressing the pedal, and the shock was never turned on except when the rat turned it on itself by stepping on the pedal.

In this box, animals invariably stepped on the pedal about 25 times during the first hour (although there was no rewarding electrical stimulation at all), because the pedal was so placed that it would be pressed when

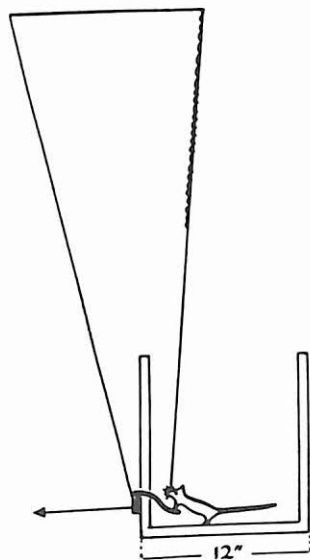


Figure 1

Diagram of apparatus by means of which a rat delivers electric shocks to its own brain. When the rat steps on the pedal, the electric circuit is closed and current is transmitted to its brain by means of implanted electrodes.

the animal looked out the only opening in the box. After the first hour there were about five responses or less per hour during hours in which no reward was produced. If the pedal-pressing produced a reward, the rate, even for the first hour, rose to 200 or more responses an hour; thus, "reward" was clearly discernible. If the pedal-pressing produced punishment, the rate dropped radically; there were only two or three responses during the total experimental procedure; thus, "punishment" was clearly discernible.

If the electrode was placed in the brain at a point at which maximum self-stimulation is produced, the rat, after its very first electric stimulation,

began to search and pursue eagerly. Its response to the first shock was to sniff in all corners of the box and manipulate quickly the objects in its path until it stepped on the pedal a second time. After the rat had pressed the pedal a second or third time, it ceased to wander and began to respond at the rate of one or two pedal-presses per second. These animals learned to press the pedal within a minute or two.

Figure 2 shows, in schematic form,

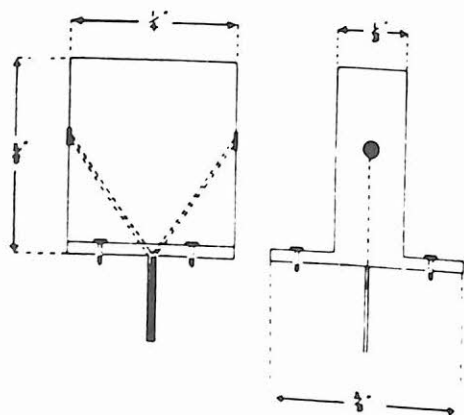


Figure 2

Diagram of plastic electrode carrier, containing a pair of silver electrodes, which is screwed to skull of rat, with wires penetrating deep into the brain, as shown in Figure 3.

the plastic electrode carrier which was screwed to the skull and a pair of silver-wire electrodes implanted in the brain through a hole in the skull. The insulated wires, which were 0.01 inch in diameter, stimulated the brain only at their tips. The electrode apparently fired cells up to a distance of at least 0.5 millimeter from its tip. Figure 3 (top) is an x-ray photograph of an electrode in place in the brain of an intact animal; in Fig. 3 (bottom) the

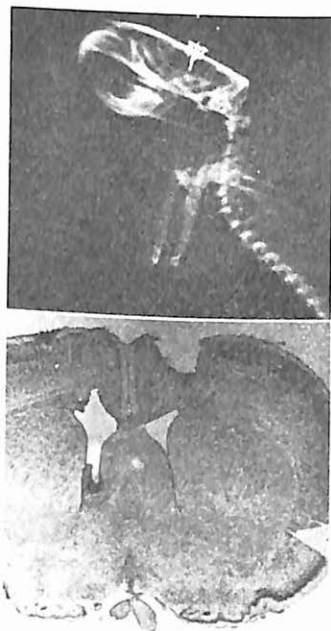


Figure 3

(Top) X-ray photograph of an electrode (deepest shaft) in place in the brain of a rat. The four shorter shafts are screws which hold the electrode carrier to the skull. (Bottom) Photomicrograph of stained brain slice from a rat sacrificed at the end of a series of experiments. The tip of the electrode track appears as darkened area. Wires are insulated along their whole length and stimulate the brain only at the tip.

tip of the electrode track appears as a blackened area on the photomicrograph of a stained brain slice. In these experiments each brain was sectioned and stained after testing; all statements about localization are based on examination of this histological material.

The stimulus was a sine-wave shock

at 60 cycles per second; the current ranged from 5 to 100 microamperes; the animal received a shock lasting for a maximum of 0.5 second. If the animal held the pedal down for less than 0.5 second, the current went off when it released the pedal. If it held the pedal down for longer than 0.5 second, the current went off automatically and the rat had to release the pedal and press it again to produce another shock.

LOCUS

Electric stimulation in most parts of the rhinencephalon, and in many parts of the hypothalamus and related structures (see cross-hatching, Fig. 4), produced the approach response (2). Stimulation in small areas in the mid-brain and in certain adjacent parts of the thalamus and hypothalamus (see stippling, Fig. 4) produced the avoidance response. Such avoidance behavior was first demonstrated in the cat, by Delgado, Roberts, and Miller (6). In the rat, the area in which stimulation produced avoidance behavior was small compared with the area in which it produced approach behavior.

The rate of self-stimulation tended to diminish steadily as the site of stimulation was moved toward the cortex. Rates as high as 7000 per hour were achieved when electric stimulation was applied in the region of the interpeduncular nucleus of the tegmentum.

With electrodes placed in the posterior hypothalamus, just in front of the mammillary body, very high rates, in the range of 5000 per hour, occurred frequently. With electrodes placed in the anterior hypothalamus, rates ranged from 400 to 1100 per hour. With electrodes placed farther forward, in the preoptic and telen-

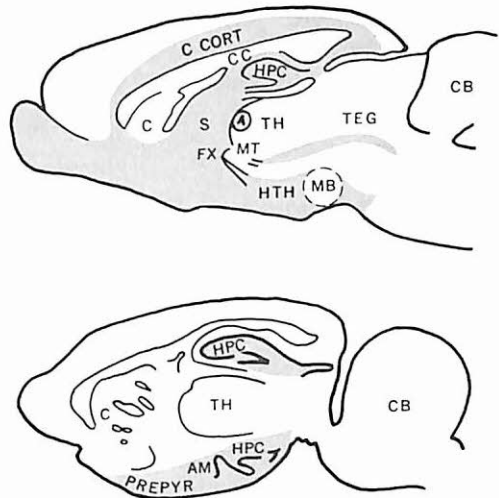


Figure 4

Medial and lateral sagittal sections of the rat brain showing, by cross-hatching, the areas where electric stimulation causes approach behavior and, by stippling, the areas where electric stimulation causes avoidance. *A*, anterior thalamus; *AM*, amygdala; *C*, caudate nucleus; *CB*, cerebellum; *CC*, corpus callosum; *C CORT*, cingulate cortex; *FX*, fornix; *HPC*, hippocampus; *HTH*, hypothalamus; *MB*, mammillary bodies; *MT*, mammillothalamic tract; *PREPYR*, prepyriform cortex; *S*, septal area; *TEG*, tegmentum; *TH*, thalamus.

cephalic areas, there was a second series of rates, ranging from quite high ones (about 3000 per hour) for the preoptic area to very low ones (about 200 per hour) for the anterior forebrain. The high rates obtained in the forebrain series were lower than the high ones of the hypothalamic series.

Thus, there was, for the hypothalamus, a decline of response rates as the electrode was moved forward. There was a similar trend for the telencephalic region, and the rates for the

whole of the telencephalic region seemed to be lower than those for the whole of the hypothalamic region.

ELECTRIC CURRENT

Studies in which the level of electric current was varied (see Fig. 5) provide some basis for explaining these differences in rate of self-stimulation. The level of the shock was raised from 0 to 150 microamperes by steps of 10. Self-stimulation rates started at chance levels of about 4 to 40 responses per hour and remained at these chance levels at 0, and sometimes at 10, microamperes. Then, usually at 20 or 30 microamperes, but sometimes at 10 microamperes or less, a threshold was

crossed and self-stimulation rates rose rapidly. As the current was raised further, the response rate showed a steady increase, undulated, or showed no further increase.

We assume that the cells and fibers excited by the electric stimulus obey the all-or-none principle and have relatively similar thresholds. If this is true, each increase in current brings cells at a greater radius from the tip of the electrode to threshold.

Thus, the steep asymptotic curves at the top of Fig. 5 indicate that stimulation of several rings of cells around the electrode tip induced the rat to stimulate itself with progressively greater frequency. The undulating curves obtained by stimulation in the

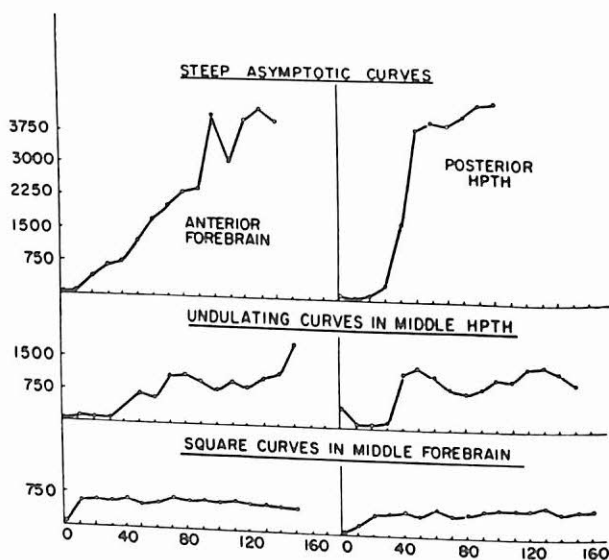


Figure 5

Electric-current functions. The hourly self-stimulation rate (plotted along the ordinate) tends to rise as current (on the abscissa) increases from 0 to 160 microamperes in steps of 10. In the middle hypothalamus the curve shows a temporary decline when the electric field invades an area where electric stimulation has a negative motivational effect. In the middle forebrain there is an abrupt rise to 500 responses an hour and then no further increase, because there is a very small field in which positive motivational effects are obtained, surrounded by a large neutral area.

middle hypothalamus suggest that stimulation of some rings of cells decreases the rate of self-stimulation; from other evidence it is known (2) that "negative" or "punishing" areas infiltrate into the "reward" system in the part of the middle hypothalamus from which these curves were obtained. The "square" curves obtained for the dorsal septal area show that only one ring of cells around the electrode tip had any motivational function; the rest were neutral. Such "square" functions have always resulted when the electrodes were implanted in or near the diagonal band of Broca; the threshold increased as the distance of the electrode from the diagonal band increased. These data on electric current level permit an important inference to be drawn: the asymptotic self-stimulation rate of an area probably depends on the number of concentric rings of "reward" cells surrounding the tip of the electrode.

The square curves indicate several other important points. Since the full rate occurred at 10 microamperes, when the electrodes were in the diagonal band, it appears that cells outside the diagonal band are neutral (the diagonal band is, thus, the only active site in the septal area). The fact that current up to 150 microamperes does not slow the rate of self-stimulation indicates that the reward units of the diagonal band are affected equally by any current from 10 microamperes (threshold) to 150 microamperes (15 times threshold). Since this function was produced repeatedly by the same rat every day for 8 months, it may be assumed that the current of 15 times threshold did no damage to the nearby cells which produce self-stimulation.

Finally, because electrodes may be placed at varying distances from the

diagonal band, it is possible to define the shape and size of the suprathreshold electric field produced by various levels of electric current.

SUMMARY

(i) The areas in which the stimulation produces the approach or rewarding effect occupy a larger proportion of the brain than do the areas in which the avoidance or punishing effect is produced. Therefore, the brain cannot be thought of as tending mainly to produce behaviors which *decrease* its own excitation, for a large portion causes behaviors which *increase* excitation. (ii) There is some sort of orderly arrangement of the rewarding effect in the rhinencephalon and related structures, with the result that response rates tend to decline as stimulation is moved forward toward the cortex; this is true both within structures and from structure to structure. (iii) Finally, by gauging the way in which rates of self-stimulation increase as the strength of the brain-shock increases, it is possible to estimate the size of the sphere, surrounding a point of stimulation, in which electric stimulation is rewarding. When the size of this sphere is large, as in the ventral posterior hypothalamus, the rate of self-stimulation at high current levels is very high. When there is only a narrow ring, as in the dorsal septal area, the rate, even at high current levels, is low.

Analytic Studies

Further studies indicate that the electrical brain-shock reward has the effect of a strong primary reward object in several different experimental situations. These studies suggest also that the electric brain shock excites cells which are normally involved in

the mediation of the effects of conventional primary reinforcers such as food and sex objects.

MAZE RUNNING

At first it seemed possible that the animal did not really seek the electric

stimulus but pressed the pedal only as a result of some sort of compulsion in a small, confined box. Maze studies (7), however, indicated that animals show day-to-day learning of a complex problem in order to get the electric reward (see Fig. 6). Test animals cer-

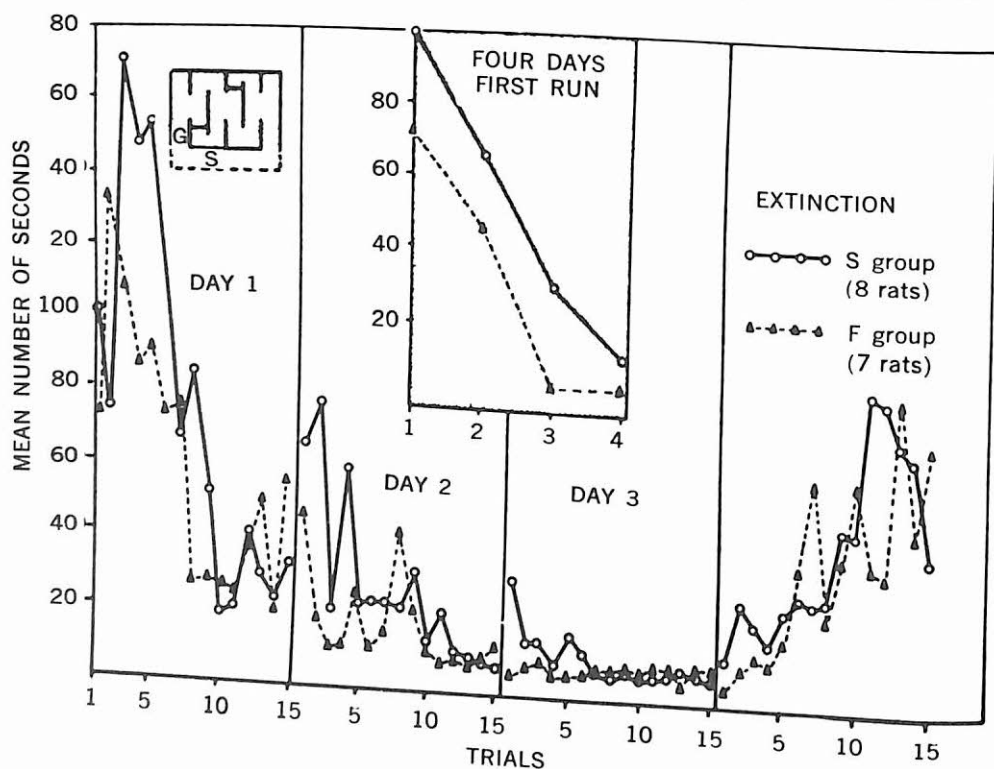


Figure 6 Maze experiment. At the upper left is a maze. The animal runs from the start position S, through the maze, to the goal position G. At G it finds a pedal, steps on it, and stimulates its brain. After the stimulation, the pedal is swung back, and the goal box becomes the start box as the animal runs the maze again. The graphs show a comparison of a group running for food as the reward (broken line) with another group running for electric shock as the reward (solid line). Both groups were 24-hours-hungry. The shock-reward group ran as fast as the food-reward group, at the end of the test, and learned almost as fast. The marked day-to-day improvement shown in the first runs of the four days is especially important. The animals of the shock-group averaged over 100 seconds on the first run the first day. They took about 50 seconds on the first run the second day, 30 seconds on the first run the third day, and 12 seconds on the fourth day. This indicates that an animal runs directly to its goal on the first run of the day, having had no pretaste of the electric stimulus to whet its desire.

tainly ran as fast as, if not faster than, others running for a food reward. In the case of a runway instead of a maze (7), animals ran much faster for the electric stimulus than for food.

OBSTRUCTION BOX

The next two studies were concerned with the strength and duration of the drive for self-stimulation. To study the strength of the drive, the obstruction box shown schematically in Fig. 7 was used. In this study (8)

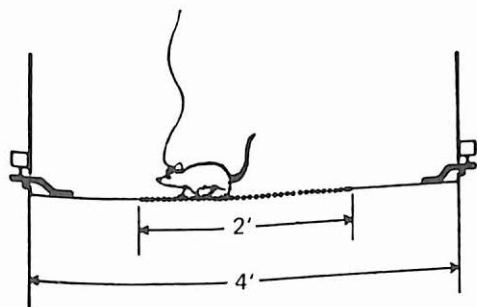


Figure 7

Obstruction box. An animal receives electric stimulations on one side, then must cross the grid to get more. The animals withstood more foot-shock for the sake of electrical reward than they did when they were 24-hours-hungry and were running for food.

the rat was permitted to stimulate itself three times at one lever. Then it had to cross a grid, which delivered an increasingly painful foot shock, to stimulate itself three more times at the opposite lever. It went back and forth until the foot shock became so great that it stopped the rat from crossing. Healthy, well-fed rats running for a brain-shock reward endured far more painful shock to the feet than did 24-hour-hungry rats running for food. The drive for self-stimulation appeared to be (in some cases) at least

twice as strong as a 24-hour hunger drive.

SATIATION?

As for the endurance of the drive for self-stimulation, two questions may be asked: (i) How many days did the drive last if rats were allowed to stimulate themselves for an hour per day? (ii) How many minutes or hours did the drive last if rats were allowed to stimulate themselves continuously?

When animals were run for periods of an hour a day, they usually maintained the same rate of self-stimulation throughout the hour and for as many days or months as they were tested. Such stable rates (shown in Figs. 8 and 9) were obtained when high supra-threshold electric current (60 to 100 microamperes) was used.

If animals with electrodes implanted in the hypothalamus were run for 24 hours or 48 hours consecutively, they continued to respond as long as physiological endurance permitted. Rats with electrodes implanted in the telencephalon, on the other hand, seemed to slow down considerably when they were shifted from a 1- to a 24-hour self-stimulation schedule. Figure 10 shows the following pattern: a rate of 2000 self-stimulations per hour maintained for 24 consecutive hours, then sleep, and then further response at a rate of 2000 self-stimulations per hour; this is typical when electrodes are implanted in the hypothalamus. Figure 11 shows rates of self-stimulation for a series of animals on such a regime. All curves obtained from animals with electrodes in the telencephalon have early, sharp inflection points. An animal with an electrode in the telencephalic septal area, which responded 1000 times per hour if it was permitted to stimulate itself for only one hour

FIFTEEN DAYS OF RESPONDING
POSTERIOR HYPOTHALAMIC ELECTRODE #315

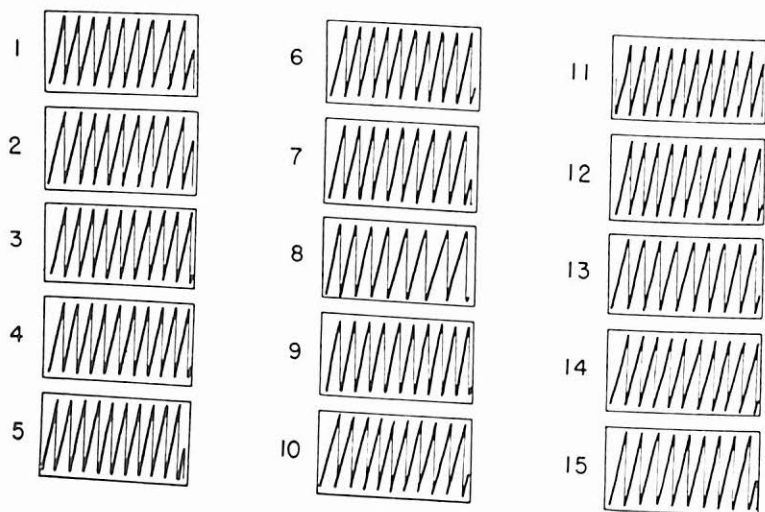


Figure 8

Response rates in maze studies of one animal for 15 consecutive days. Each square indicates a 1-hour test period on a given day. Each time the record goes from the bottom to the top of a square, the animal has made 500 responses; thus, each peak represents 500 responses or self-stimulations (9). The slope of the rising edge of each "sawtooth" indicates the response rate. There are about ten peaks in each day's record; thus, the animal was responding at the rate of about 5000 responses an hour. The electrode in this case was implanted in the posterior hypothalamus. The same rate is maintained all through the hour, and it is also maintained on all 15 days. Such rates are maintained, in fact, for several months in a row—that is, for as long as we choose to run the animal.

per day, responded only 1000 times in 24 hours if it was permitted to stimulate itself for 24 hours per day. Thus, animals with electrodes in the telencephalon appeared to show some genuine satiation. No similar satiation appeared in animals with electrodes in the hypothalamus.

EFFECTS OF DRIVES

Further studies showed that the animal's taste for the electric reward is often sensitive to basic drives in the way that its taste for conventional re-

wards is. The extensive reward system appears to break down into subsystems subservient to the different basic drives; there appears to be a food-reward system, a sex-reward system, and so on.

If electric stimulation at some points fires cells that mediate food reward, the animal's appetite for self-stimulation at this point may go up and down with hunger as its appetite for food does. A large series of animals has been tested in order to compare their rates of self-stimulation when hungry with their rates when full. When tests

FIFTEEN DAYS OF RESPONDING
FOREBRAIN ELECTRODE # 215

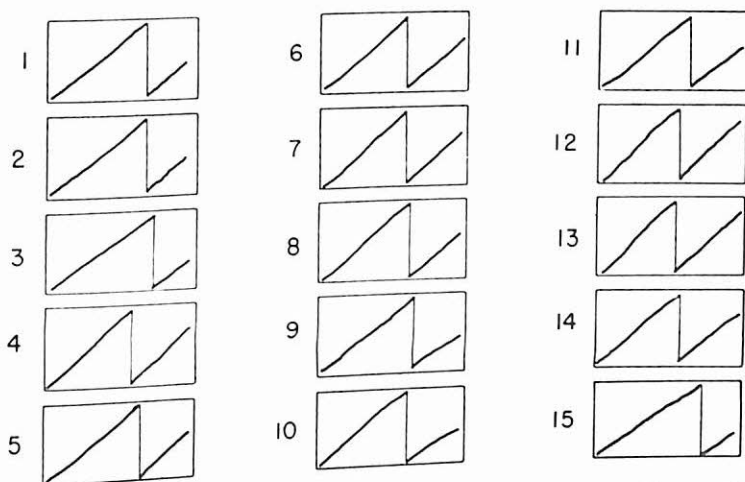


Figure 9

Diagram similar to Figure 8, showing response rates in maze studies of an animal with an electrode implanted below the septal region. Here the rate is about 750 self-stimulations an hour.

were made at a constant current of 65 microamperes with a set of electrodes placed in the midline of the brain, in the ventromedial hypothalamus, and in the septal area, hunger seemed to have an important positive effect, increasing self-stimulation rates (9, 10). When, however, animals were tested at a series of current levels, a somewhat different picture of the hunger system appeared.

In the latter experiments (as shown in Fig. 12) animals were tested every day at a series of shock levels; accordingly, they yielded a series of response rates (usually of increasing magnitude). Each shock level was maintained for 8 minutes, response output was recorded, and then the shock was shifted to the next higher level. Zero microamperes (no reward) was usually maintained for the first interval, and then the shock level was raised to 5

microamperes, 10 microamperes, and so on. Response rates started at very low (chance) levels during the first interval and gradually rose to an asymptote characteristic of the area of stimulation.

Animals were run alternately—one day hungry and the next day full—to see whether this would change the rate of self-stimulation during the various intervals. Many animals responded faster when hungry and slower when sated, but this difference appeared only at a limited range of electric shock levels. The shock level at which the hunger difference appeared is called the "threshold of the hunger effect."

Before discussing these thresholds I shall mention the gross findings of this recent hunger study. When electrodes were in the hypothalamus, all large hunger effects on self-stimulation were

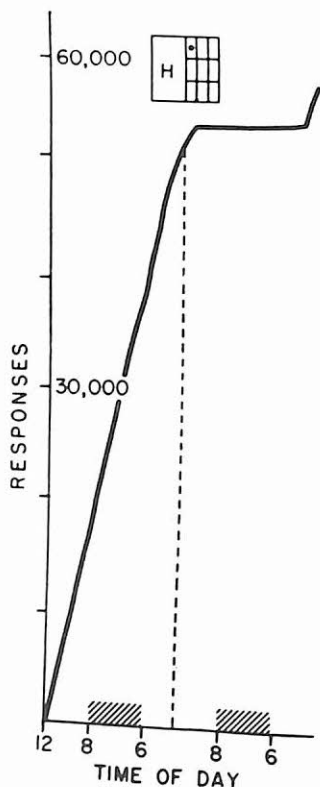


Figure 10

Graph of continuous self-stimulation for a 48-hour period. Cumulative response totals are plotted along the ordinate, hours along the abscissa. The experiment started at noon; cross-hatching indicates darkness from 8 P.M. to 6 A.M. The animal (with an electrode implanted in the anterior medial hypothalamus) stimulated itself at a rate of more than 2000 responses an hour for 26 hours, then slept, and then resumed self-stimulation at the same rate.

obtained with those electrodes which were clustered in the posterior sector. No hunger effects were found when electrodes were in the portion of the anterior hypothalamus that was ex-

plored. With electrodes in the telencephalon, however, strong hunger effects appeared again in a rather unclear pattern. With electrodes in the posterior hypothalamus, the interesting and orderly picture shown in Fig. 13 was obtained.

Thresholds for the hunger effect declined as the location of the electrodes approached a point about 1.25 millimeters lateral to the midline at a point just in front of the mammillary body. If we assume that a hunger-reward center is located at this point, we find that it can be reached by 10 microamperes of current from electrodes at points about 0.25 millimeter away and by 20 microamperes from electrodes at points about 0.5 millimeter away. From these data we may surmise (see Fig. 14) that a 50-microampere stimulus has a suprathreshold field with a diameter of about 1 millimeter.

Electric stimulation in the whole posterior hypothalamus has rewarding effects, for stimulation through electrodes anywhere in this area produces very high rates of response even at the lowest current levels. Only stimulation of areas shown schematically at the lower left of Fig. 14, however, seems related to hunger; the other areas must be controlled by other drives.

A hunger differential appears, we may surmise, whenever an electric field is increased to the point where there is a hunger-sensitive area on the boundary between the supra- and subthreshold parts. If this is true, our data suggest a sharp localization of a hunger reward center in the hypothalamus and a very compact relation between it and the other drive-reward centers in the hypothalamus—a relation so compact that different effects will be achieved at different current levels.

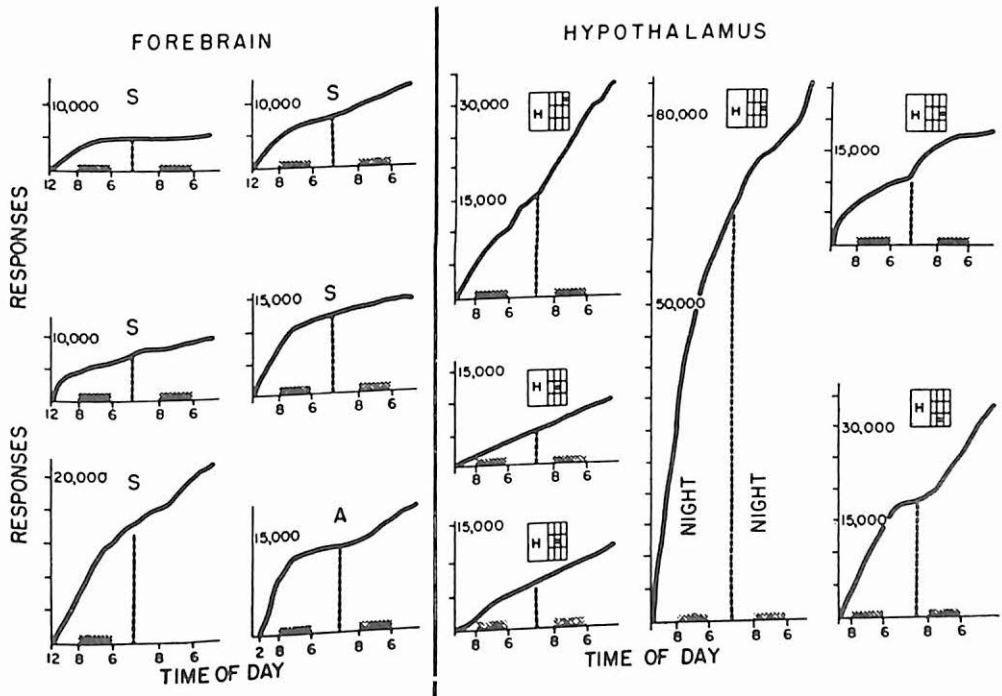


Figure 11

Graphs of comparative responses of animals with electrodes implanted in forebrain and of others that had electrodes in hypothalamus, in a continuous 48-hour experiment. Animals with electrodes in the forebrain show a point of satiation which does not appear in animals with electrodes in the hypothalamus.

Androgen-Level Studies

In androgen-level studies (11), animals were castrated after they had been trained to press the lever for brain shock. After castration, rates of self-stimulation were measured for 14 days of declining androgen level. Then testosterone propionate was injected, in doses of from 1 to 5 milligrams per animal, and rates were measured over a period of days while androgen levels rose and fell again. A series of testosterone injections was also given to see how rates of self-stimulation were maintained under androgen replacement therapy.

Two sets of data from this study are shown in Figs. 15 and 16. First, with an electrode in the dorsomedial caudate nucleus, an all-or-none relation between rate of response and testosterone level was obtained. The animal was tested each day at a series of levels of current, from 15 to 55 microamperes. The animal responded well for a 15-microampere current when androgen levels were high; it did not respond at all, even for a 55-microampere current, when androgen levels were low (see Fig. 15). The animal was taken through several androgen cycles. At high levels it always responded; at low levels it did not. At

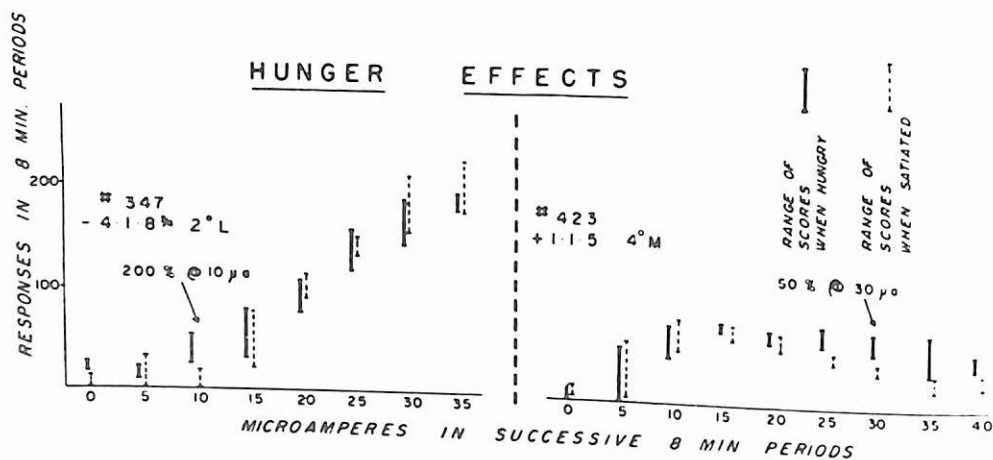


Figure 12

Hunger effects. Response output is plotted against electric current; each bar or dotted line represents a range of scores obtained on several days. The testing started with current at zero. Current was increased after 8 minutes to 5 microamperes, then to 10 microamperes, and so forth, being maintained for 8 minutes at each level up to 35 or 40 microamperes. Response rates gradually increased, approaching an asymptote characteristic of the area of stimulation. The heavy black lines represent response ranges for three different days when animals were 24-hours-hungry; the broken lines represent ranges for three days when the animals were full. At the left, a sharp hunger differentiation at 10 microamperes is indicated. There is no overlap of scores at this point. The animal averaged 200 percent more responses when it was hungry. However, at all other levels of current this difference disappears. At the right are scores for a different animal with a different electrode placement. Here there is a marked separation of scores at the higher current levels of 25, 30, and 35 microamperes. But at the lower levels, where the animal had quite high response rates, there is no difference between scores while it is hungry and scores while it is full.

the termination of its career the rat was responding at a very high level, and post-mortem examination showed that its seminal vesicles were not only as large as those of noncastrated rats but, indeed, were larger than normal. Thus, it is possible to get all-or-nothing drive effects from self-stimulation with electrodes placed in some parts of the caudate nucleus.

The second point is that there is an inverse relation between androgen effects and hunger effects on the rate of self-stimulation (see Fig. 16). Hunger

effects and androgen effects on rate of self-stimulation were measured in 16 animals, with electrodes in different parts of the hypothalamus and telencephalon, on 1-hour self-stimulation runs with a stimulus of 1 volt (about 60 to 80 microamperes). Animals which showed a strong increase in rate of self-stimulation in response to androgen showed a decline in rate of self-stimulation in response to hunger, and vice versa. In other words, if hunger raised the response rate, androgen lowered it; if androgen raised the re-

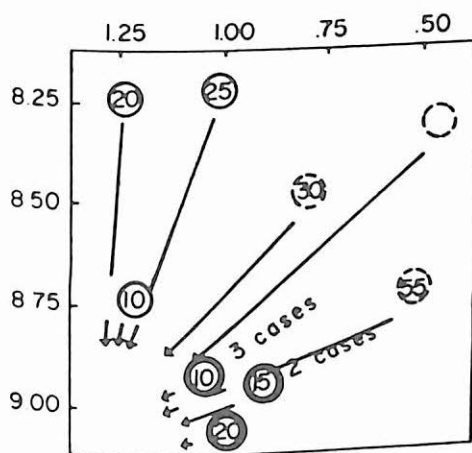


Figure 13

Hunger map. This is a schematic map of the region of the posterior hypothalamus. (Electrodes 8 to 9 mm long are implanted 4 mm behind the primary skull marking 4 mm behind the primary skull marking (in bregma.) The ordinate gives distance (in millimeters) from the top of the skull. The abscissa gives distance (in millimeters) from the midline. Thus, the right side of the drawing represents the medial portion of the hypothalamus. The very dark circles indicate electrode implantations accompanied by very large hunger effects; plain circles indicate those accompanied by moderate effects; broken circles indicate those accompanied by very small effects. The broken circle without a numeral indicates no hunger effect at all. The numerals inside the other circles indicate the level of current (in microamperes) at which the hunger effect could be detected. In the case of the electrodes represented by the cluster of circles at the lower left-hand corner of the map, levels of current of about 10 microamperes produced the hunger effect. With electrodes represented by circles above this point, current levels of 20 and 25 microamperes gave the hunger effect. With electrodes nearer the midline, currents of 30, then of 55, microamperes were required to get

sponse rate, hunger lowered it. If we presume that the placement of the electrode determines the degree of sensitivity to hunger or to androgens, this study gives a firm basis for expecting to find anatomical differentiation between the hunger-reward system and the androgen-reward system.

Other experiments showed the effect of forced electric stimulation on food consumption (consummatory behavior). In these experiments (12), electrodes were implanted and were used first to test the effect of stimulation on food consumption. The animals were maintained on an *ad libitum* diet. Eating, under stimulation, in stimulation boxes, was measured for 1 hour per day with current of about 80 microamperes applied for 0.5 second every 10 seconds. Other groups were stimulated with current of 25 microamperes applied for 0.5 second every second. Measures of food intake under stimulation were matched against food intake during control weeks during which there was no stimulation. As has been reported by others (13), stimulation in the ventral posterior hypothalamus at points about 1.5 millimeters lateral to the midline caused an increase in eating; stimulation medial to these points sometimes, but not always, caused a decrease in eating. After these tests were completed, the same animals were subjected to self-stimulation tests

hunger differentiation, and even then it was small. With the electrodes represented by circles in the top medial part of the square, no hunger effect is detected at any of the levels of stimulation used. It is almost, but not quite, true that the distance from a point 9 mm deep and 1.25 mm lateral determines the amount of current needed to get the hunger effect.

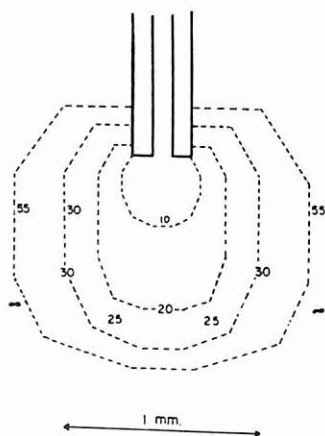


Figure 14

Electric fields for different levels of current. A current of 10 microamperes seems to produce a suprathreshold electric field which extends more than 0.25 mm downward but a much shorter distance out to the sides and a still shorter distance upward; 20 microamperes seems to reach about 0.75 mm downward; 55 microamperes, about 1 mm. If reference is made to this schematic drawing of the field in connection with each point on the diagram of hunger effects (Figure 13), it may be seen that the boundary of the suprathreshold field cuts through the supposed hunger point at the shock level where the hunger difference appears. This drawing is intended to suggest orders of magnitude but not exact relationships.

with the same levels of current. The lateral electrode placements, in areas where stimulation seemed to increase hunger drive, were the ones that usually produced extremely high rates of self-stimulation. The medial placements, in areas where stimulation seemed to reduce the hunger drive, ordinarily produced much slower rates of self-stimulation.

To avoid the notion that there is a food-reward "center" in the posterior

hypothalamus, it should be mentioned that a similar relationship between hunger and electric stimulation seems to exist in several more anterior parts of the brain. Figure 17 shows the effect from a rewarding electrode placement in the telencephalon. Electric stimulation at this point increased eating behavior by almost 50 percent, and the animal stimulated itself at a rate of over 1000 responses per hour with no change in level of current.

There is a superficial anachronism in these data correlating drive increase (measured by eating behavior) and reward (measured by self-stimulation); reward has classically been thought to be correlated with drive reduction. There is, of course, the possibility that some conflicting drive is reduced by the stimulation, thereby permitting the animal to eat. There may also be neighboring drive and reward areas, both activated by the same electric field. However, the correlation of rewarding properties with a stimulus which produces consummatory behavior need not be surprising. If we think, for example, of a stimulus to a sexual consummatory response, we might expect some rewarding properties. Why not expect a similar result in the case of hunger?

Thus, our analysis permits the following generalizations: (i) the electrical reward is effective where more conventional rewards are effective—in a Skinner box, runway, complicated maze, and obstruction box; (ii) the reward has the same effect from day to day over long periods of time; (iii) with stimulation through electrodes in some brain areas, hunger augments the rate of self-stimulation, and with stimulation through electrodes in other brain areas, androgens augment the rate; (iv) the rewarding stimulus often appears to produce a temporary

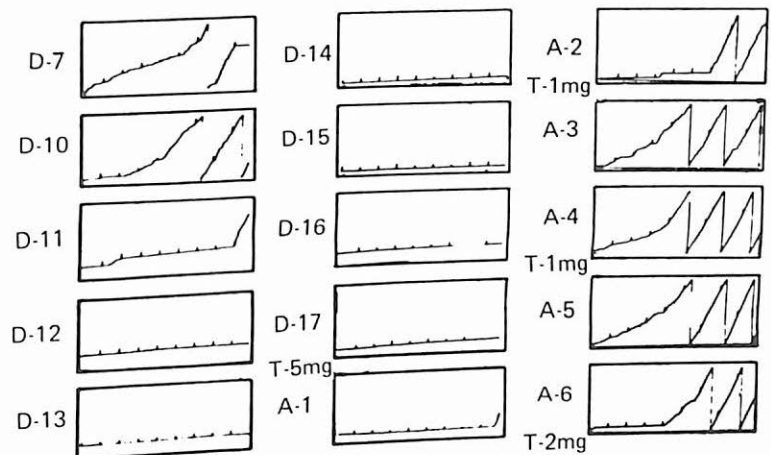


Figure 15

Effect of androgen on threshold of response to stimulation, showing threshold changes from day 7 after castration to day 17 after castration and reversal of these changes by injection of 5 mg of testosterone propionate in oil. Animals are tested each day at levels of current that increase from 15 to 55 microamperes by steps of 5 microamperes.

increment in some consummatory behavior.

Drug Studies

The last part of this article deals with more practical considerations. It seems that certain behavior disorders might be benefited if "reward" or "pleasure" systems of the brain could be selectively controlled by use of pharmacological agents. It might be important to control one of the drive-reward systems without affecting the other systems, or to control the forward or cortical reward systems without affecting the posterior, hypothalamic systems.

For these reasons we have hoped to find differential sensitivity of different reward placements to different drugs with known emotional or psychological effects. In recent drug studies (14), series of levels of current were

used, as in the drive studies. Animals were allowed to stimulate themselves for eight minutes at each level of current, starting at zero and working up to 40 or 50 microamperes. As was indicated previously, the animals did not respond at zero and 5 microamperes. At 10, 15, or 20 microamperes, the animals began to respond. The eight-minute interval in which they began to respond was called the threshold interval. A drug was injected, to become effective at the beginning of the threshold interval (which had been determined by several weeks of previous control testing).

TRANQUILIZERS

In the bar graphs of Fig. 18 the rate of self-stimulation is shown for successive levels of current after threshold has been reached (these graphs

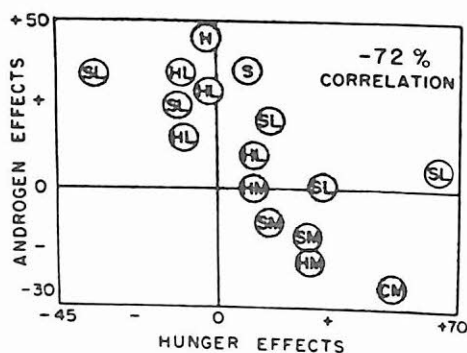


Figure 16

Relation of hunger effects and androgen effects. Along the abscissa, differences from control rates caused by 24 hours of hunger are plotted as percentages of control rates. Along the ordinate are plotted similar differences caused by administration of 2 milligrams of testosterone to castrated animals. Each point represents one electrode. *SL*, lateral septal area; *HL*, lateral hypothalamic area; *SM*, medial septal area; *HM*, medial hypothalamic area; *CM*, medial caudate area. Animals were run at high constant levels of electric current. Stimulation points positively affected by hunger tend to be negatively affected by androgens, and vice versa.

start, that is, with the threshold interval). Rectangular outlines show the average rate of response for the day before, and the day after, the drug test. The dark bars indicate the rate of self-stimulation when the animal was under the influence of the drug. The dotted lines show the degree of the effect of the drug. In Fig. 18 the data are for chlorpromazine injected intraperitoneally at 2 milligrams per kilogram. In the case illustrated at the left of Fig. 18, in which the electrode was in the middle hypothalamus, none of the self-stimulating response was eliminated at threshold, only 25 per-

cent of the response was eliminated at the next level, and 21 percent was eliminated at the third level, after which all inhibitory effect of chlorpromazine was over, in this case. In the case illustrated at the right, in which the electrode was placed in the ventral posterior hypothalamus, the effect was much more pronounced; self-stimulation is totally eliminated in all six of the intervals shown.

A map of the effects of chlorpromazine obtained with 31 cases has been made. It shows that the effects of chlorpromazine are strongly inhibitory on self-stimulation when electrodes are placed in the ventral posterior hypothalamus. With electrodes in the middle hypothalamus, there is a somewhat less marked effect, and with electrodes in the anterior hypothalamus, there is very little effect. The effects are small with electrodes in the anterior preoptic region but seem to become quite strong again with electrodes in parts of the septal region.

With electrodes in the posterior hypothalamus and in the anterior hypothalamus, the same drug has strikingly different effects on the rate of self-stimulation. Since the drug does not greatly slow self-stimulation via electrodes implanted in the anterior hypothalamus, there is some assurance that the drug does not impede behavior as such. Since the drug does greatly slow self-stimulation via electrodes implanted in the ventral posterior hypothalamus, there is some assurance that the drug acts selectively, either on this part of the brain or on some of the areas to which it projects, to produce its rewarding effect. Furthermore, it is reasonable to assume that the areas sensitive to the drug are not essential to all rewarding effects, because anterior hypothalamic stimulation has its rewarding effect in spite of the drug.

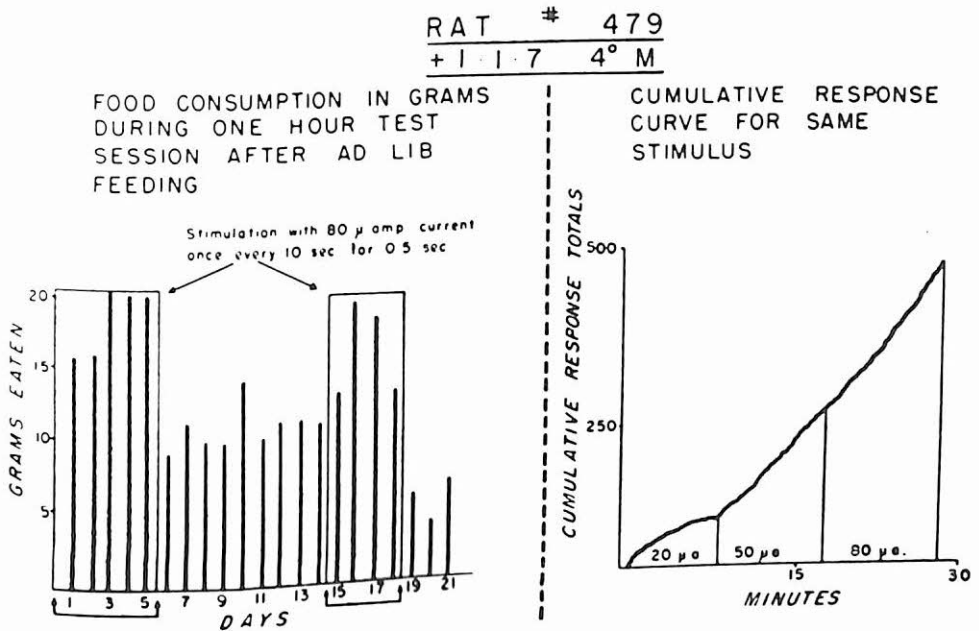


Figure 17

Eating behavior produced by "rewarding" stimulus.

PSYCHOTOMIMETICS

Another example of a differential drug effect is the interaction of serotonin and lysergic acid diethylamide. Serotonin is supposedly a transmitter substance in the brain, and lysergic acid diethylamide is a psychotomimetic (producing effects somewhat similar to psychotic hallucinations). Studies by other workers (15) have suggested that in some cases serotonin might be antagonistic to the effects of lysergic acid diethylamide. When lysergic acid diethylamide was injected intraperitoneally, just prior to the threshold interval, there was usually an inhibitory effect on self-stimulation. In most cases the drug was effective in both the first and second intervals after injection. The effect of the drug appeared most dependably in the second eight-minute interval after injection;

therefore, we studied the inhibitory effect on self-stimulation in this second interval in further analysis. We found two types of interaction between lysergic acid diethylamide and serotonin. In the first, lysergic acid diethylamide alone had a strong inhibitory effect on self-stimulation; when serotonin was administered half an hour before the lysergic acid diethylamide injection, however, there was no effect of the latter at all. Finally, bromo-lysergic acid diethylamide which is like lysergic acid diethylamide but is supposed not to cross easily from the blood into the brain, had no inhibitory effect on self-stimulation. In the second case, lysergic acid diethylamide had a strong effect again. This time, however, serotonin did not antagonize the effect, and bromo-lysergic acid diethylamide had the same effect as lysergic acid diethylamide itself.

CHLORPROMAZINE EFFECTS

(2 mg / Kg)

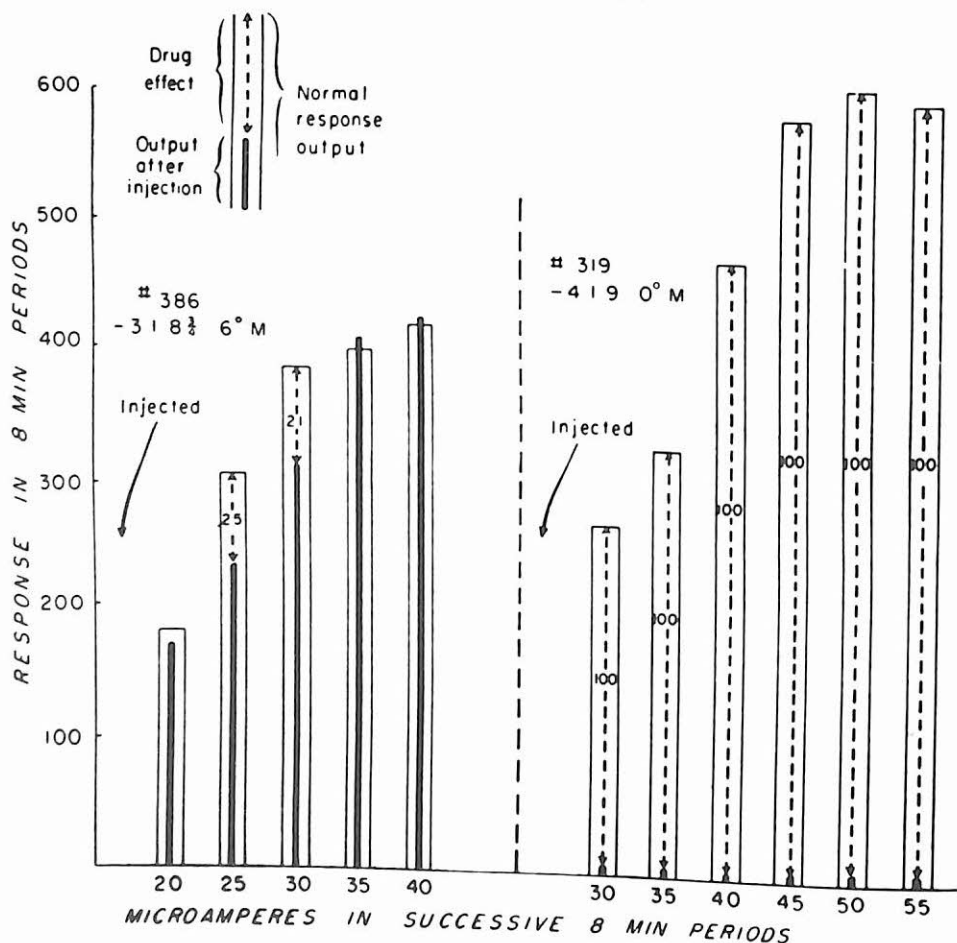


Figure 18

Differential effects of chlorpromazine. Rectangles show normal response output. Black bars show output under influence of chlorpromazine. The animal for which data are given at left was almost unaffected by the drug. The animal for which data are given at right was almost totally inhibited; the difference is attributed to the more posterior position of the electrode in the latter case.

The two effects on self-stimulation are mapped in Fig. 19. We see that when electrodes are in certain delimited parts of the hypothalamus and telencephalon, serotonin fails to antagonize the effects of lysergic acid

diethylamide, and in these cases bromo-lysergic acid diethylamide has the same effects as lysergic acid diethylamide. With electrodes implanted in other clear-cut regions, serotonin does antagonize the effects of lysergic

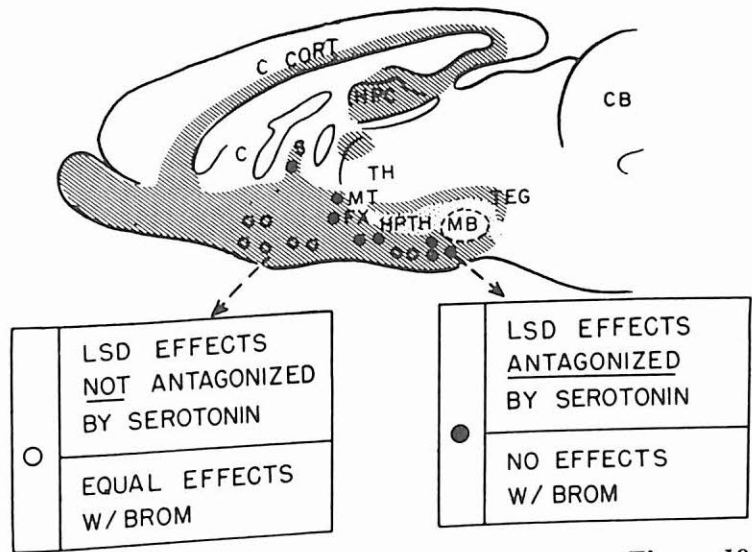


Figure 19

Map showing points where effects of lysergic acid diethylamide on self-stimulation are antagonized by serotonin and other points where the effects are not antagonized. At the points where serotonin fails to protect against lysergic acid diethylamide, bromo-lysergic acid diethylamide has the same effects as lysergic acid diethylamide itself.

acid diethylamide, and these are the cases in which bromo-lysergic acid diethylamide has no inhibitory effects on self-stimulation. Here again, we have evidence of chemical differentiation between different parts of the reward system.

In some earlier tests on chlorpromazine and reserpine, a constant and quite high voltage level was maintained (16). Striking inhibition of response tendencies by reserpine was found when electrodes were implanted in the posterior hypothalamus, but no effects were found when electrodes were in the telencephalic region. Later work with graduated levels of electric current did not show such a sharp differentiation. The inhibitory effects of reserpine on self-stimulation in the hypothalamus do not appear at low voltages but only at very high ones.

These data indicate that reserpine may selectively lower seizure thresholds for stimulation via the posterior hypothalamus. Petit mal states, elicited by stimulation at high voltages plus the reserpine, seem to account for the seeming inhibitory effect of reserpine on self-stimulation via the hypothalamus.

Summary

My conclusions are these: (i) The cells which mediate primary rewarding effects are located in a midline system running from the midbrain through the hypothalamus and midline thalamus and into the subcortical and cortical groups of the rhinencephalon. (ii) The cell groups which mediate primary rewarding effects are different from those which mediate

primary punishing effects. (iii) Despite this relative independence, there are, undoubtedly, relationships of mutual inhibition existing between these two systems. Rewards do, among other things, tend to reduce sensitivity to pain, and punishments do tend to reduce rewarding effects. (iv) These primary reward systems of the brain are subdivided into specific drive-reward subsystems mediating the specific drives such as hunger and sex. (v) Because there are also subsystems of this set of rewarding structures sensitive to different chemical effects, it is reasonable to hope that eventually it will be possible to control the reward systems pharmacologically in cases where behavior disorders seem to result from deficits or surfeits of positive motivation.

References and Notes

1. These studies were aided by a grant from the Foundations Fund for Research in Psychiatry ("Brain Mechanisms and Motivation") to H. W. Magoun and J. Olds, by contract NR 144-102 [Nonr-233 (32)] between the Office of Naval Research, U.S. Department of the Navy, and the University of California, and by a grant from the Carnegie Corporation of New York ("Brain Organizations and Behavior") to R. B. Livingston and D. B. Lindsay.
2. OLDS, J., *J. Comp. and Physiol. Psychol.*, **49**, 281 (1956).
3. OLDS, J. and MILNER, P., *ibid.* **47**, 419 (1954).
4. HESS, W. R., *Das Zwischenhirn* (Schwabe, Basel, Switzerland, 1949).
5. SKINNER, B. F., *The Behavior of Organisms* (Appleton-Century, New York, 1938).
6. DELGADO, J. M. R., ROBERTS, W. W., MILLER, N. E., *Am. J. Physiol.* **179**, 587 (1954).
7. OLDS, J., *J. Comp. and Physiol. Psychol.*, **49**, 507 (1956).
8. This study was made in collaboration with J. C. Sinclair.
9. OLDS, J., "Differential Effects of Drives and Drugs on Self-Stimulation at Different Brain Sites," in D. E. Sheer, Ed., "Electrical Stimulation of the Brain," proceedings of the University of Houston and the Hogg Foundation Symposium on Brain Stimulation, 26-28 Apr. 1956, in press.
10. ———, "Neurophysiology of Drive," *Psychiat. Research Repts.* **6**, 15 (1956).
11. These studies were made in collaboration with B. V. Critchlow.
12. These experiments were made in collaboration with R. Wendt.
13. DELGADO, J. M. R. and ANAND, B. K., *Am. J. Physiol.* **172**, 162 (1953).
14. These studies were made in collaboration with K. F. Killam and S. Eiduson.
15. WOOLEY, D. W., *Proc. Natl. Acad. Sci. (U.S.)* **41**, 338 (1955).
16. OLDS, J., KILLAM, K. F., BACH-Y-RITA, P., *Science* **124**, 265 (1956).

INTRODUCTION

Some of the most fascinating data in neuropsychology has come from studying patients who have undergone accidental injury of the brain. In the following article Professor Teuber reviews how psychologists study the specific function of various areas such as the frontal lobe. In patients where this part of the brain has been damaged or removed for any of a variety of reasons, psychological tests have shown that it possesses many specific functions in man, and houses many of man's particular behavioral patterns. All of us have probably known someone who has had brain damage of some kind. After reading this paper you should be able to assess how the neuropsychologist proceeds in studying brain-damaged patients, and how you yourself might be able to detect particular brain damage in an individual by observing his behavior.

THE RIDDLE OF FRONTAL LOBE FUNCTION IN MAN

Hans-Lukas Teuber

Introduction

This paper records a recurrent perplexity and a recent attempt at its resolution. Let me begin with the perplexity. Man's frontal lobes

have always presented problems that seemed to exceed those encountered in studying other regions of his brain. In striving to give a functional interpretation of anatomic pattern, we have therefore tended, in our laboratory, to begin by considering the occipital lobes and their processing of visual information (Teuber & Bender, 1949; Teuber et al., 1960), and to progress from there to analysis of pa-

Reprinted from The Frontal Granular Cortex and Behavior by J. Warren and K. Akert. Copyright © 1964, McGraw-Hill, Inc. Used by permission of McGraw-Hill Book Company.

rietal and temporal lobe syndromes involving touch and hearing and higher processes (Teuber et al., 1951; Semmes et al., 1960). In this way, one can spend one's life as an investigator without ever crossing the central fissure. And well one might—because it appears that by starting at the posterior end of the brain, it may become impossible in principle to reach the frontal lobes.

I now believe that analysis of frontal lobe functions requires that we begin from that end, considering how these anterior parts act upon posterior, and predominantly sensory structures. What is needed is a 180° turn (Teuber, 1961b). It is hardly an accident that the frontal lobes have continued to pose such insuperable difficulties to interpretations in traditional stimulus-response terms. Such conventional attacks on the problem go from sensory to motor functions; if the opposite approach is required—going from motor to sensory, as we shall suggest—then it may become understandable why frontal symptoms have been so perplexing.

Comparison with Infrahuman Forms

This may be as true for our approach to frontal lobe function in monkeys and carnivores as it seems to be true for man. During the last three days two things have become abundantly clear: First, despite their family resemblance, these assorted aspects of frontal lobe dysfunction, in infrahuman species, are dissociable, since any one of these symptoms can be shown to occur without the others. This dissociation can be demonstrated either by making sufficiently selective lesions (e.g., frontolateral versus or-

bital), or by considering different species.

The second point, more painful to admit, is this: Despite the conspicuous efforts and ingenuity that went into behavioral analysis of these frontal lobe symptoms, their interpretation, in terms of some basic alterations in carnivore or simian behavior, remains elusive. We still do not have a satisfactory functional interpretation of the delayed-response task (Carr, cited by Hunter, 1913; Hunter, 1913; Jacobsen, 1935)—to take just one example—since we have been shown by Mishkin (Chapter 11) that the delayed-response deficit cannot be reduced to an inability to overcome response preferences (Mishkin et al., 1962). Moreover, none of us seems ready to say how it comes about that the classical frontal lobe symptoms fail to appear if the lesions are placed quite early in an animal's life. Clearly, functional interpretation will require new ways of looking at the symptoms. I am hopeful that such new viewpoints might be derived eventually from continued study of frontal lobe syndromes in man.

Contradictions in the Clinical Literature

To raise such hopes at this stage may seem unreasonable in view of the bewildering variety of symptoms that have been reported as consequences of frontal lesions in man and the still greater variety of interpretations in terms of functions. There certainly is no other cerebral structure in which lesions can produce such a wide range of symptoms, from cruel alterations in character to mild changes in mood that seem to become undetectable in a year or two after the lesion. Lack of

precise knowledge of the lesion, in the majority of clinical cases, might account for the difference from the situation found in animal experiments where there is such a dependable relationship between certain fractional frontal lesions in the monkey (bilateral sulcus principalis removals, or destruction of caudate nucleus) and a specific alteration in behavior (delayed-response deficit). Linked to the uncertainty about symptoms in man is the bewildering range of interpretations. These views go from the extreme position that the frontal lobes in man are his highest centers, or "organs of civilization," to the assertion that large uncomplicated removals of prefrontal cortex need not be followed by any behavioral deficits.

Sources of Contradictions

There are probably a number of reasons for this unsatisfactory state of affairs: the radical nature of earlier claims; the pitfalls in dealing with the single case; the hidden bias in many studies of groups of cases; and, finally, as in the animal experiments, the lack of a suitable theory. We shall briefly consider each of these points, then turn to a sketch of a theory; after that, we shall present our own experimental results in the light of this tentative theoretical approach.

Extreme Views. Discussion of frontal lobe syndromes in man have been marred by an extravagance of claims not found for other structures. During the first third of this century, there developed a substantial majority opinion which ascribed to the frontal lobes all the highest conceivable functions—from "abstract behavior," foresight, and intellectual synthesis, to capacity for ethical behavior, control of effect,

awareness of self, and recent memory (Ackerly, 1935; Brickner, 1936; Goldstein, 1936; Goldstein & Katz, 1937; Goldstein & Scheerer, 1941; Rylander, 1939, 1943). In the early forties and through the fifties of the century, the pendulum swung in the other direction—toward a denial of any especial importance of human prefrontal structures (Hebb, 1939, 1945; Mettler, 1949; Pollack, 1955; Teuber, 1959; Teuber et al., 1951). Both assertions—that of a crucial importance of frontal regions, and that of their being dispensable—have often been defended by reliance on the single case (Ackerly, 1935; Beringer, 1941; Brickner, 1936).

Single Cases. The famous patient, "Joe A.," of Brickner (1934, 1936, 1939, 1952), with his seriously deteriorated behavior, contrasts with the case of K.M., reported by Hebb & Penfield (1940), where bilateral removal of frontal tissue (including an epileptogenic scar) seemed to restore normal behavior in a young man who had suffered a depressed bifrontal fracture at the age of sixteen. Brickner's case, incidentally, turned out, on autopsy, to harbor multiple meningiomas besides the one that had been removed originally together with his frontal poles (Brickner, 1952). One of these remaining tumors was the size of an apple, and was found in the parieto-occipital lobe substance.

It is also obvious that frontal tumors should be, on the whole, more devastating than trauma—even if the tumor is limited to prefrontal regions. The disposition of blood vessels adds to this complexity (Critchley, 1930); we need only to remind ourselves of the role of the anterior cerebral artery (Figure 20.1), which supplies not only frontal-cortical, but also deep sub-

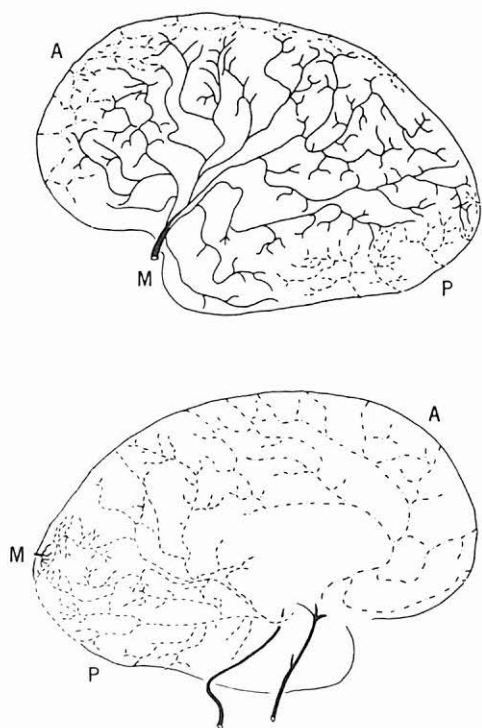


Figure 20.1

Distribution of cerebral arteries (A, anterior, M, middle, P, posterior). Note vast extent of region supplied by branches of the anterior cerebral artery. The lateral aspect of the brain is shown above, the medial below. A morbid process compromising the anterior cerebral artery produces disturbance beyond the confines of the frontal lobe, particularly in medially situated structures. (From Teuber et al., 1960.)

cortical structures, and reaches back into central and postcentral regions.

Even less is known about the role of age in determining the outcome of frontal lesions. Some of the strikingly "negative" cases seem to have sustained their frontal lesions quite early in life (Hebb, 1945), in the first or second decade. Other equally isolated

case reports (such as those by Ritchie Russell, 1959; or Ackerly, 1935) point in the opposite direction, a seriously disabling effect of bifrontal lesions incurred in childhood.

Studies of Groups with Frontal Tumors. The contradictions between studies carried out in the past on groups of patients is no less puzzling. Maximal deficits were found after frontal tumors, as compared with tumors elsewhere in the brain, both by Rylander (1939, 1943) and by Halstead (1947); yet these results are contradicted by the work of Pollack (1955) and Battersby et al. (1955) and that of others who used similar tests and tried to assemble similar cases. Perhaps an important source of difficulty lies in the fact that Rylander selected his cases in such a way as to produce, unwittingly, a greater severity of deficit in the frontal group, merely by excluding cases of aphasia, apraxia, and agnosia from his total sample; these clinical changes are more likely to appear with serious posterior lesions.

Generally, the larger size of the frontal lobes and their comparative distance from specific motor or sensory or language areas will allow a tumor or abscess to grow in the frontal areas for much longer periods without being detected. Thus, "unselected" clinical groups—no matter how numerous—may reflect a curious bias in favor of larger lesions in the frontal lobe cases, and smaller lesions with involvement of other parts of the brain.

Studies of Groups of Cases after "Psychosurgery." Work on cases of surgical lesions, as in lobotomies (Partridge, 1950; Petrie, 1952a) and topectomies (Mettler, 1949, 1952), has difficulties of another, but equally serious kind. There are no contrasting groups with lesions outside the frontal lobes, so that we have no way of tell-

ing whether any symptoms produced by the frontal lesions are in fact specific for frontal pathology, or are to be expected equally with lesions in other lobes. Still more serious is the fact that the surgical trauma is superimposed in most cases of psychosurgery on preexisting psychiatric illness, or intractable pain states. Thus, failure to find effects in particular dimensions of behavior after psychosurgery might merely indicate that the pathology produced by the surgeon's knife is submerged in the preexisting abnormalities. This point has been made most forcefully by Rosvold & Mishkin (1950), who showed that test intelligence after lobotomy did not drop in relation to scores obtained just prior to operation, and thus during the psychiatric disease that led to it. Yet both types of scores were lower than those obtained much earlier, before the onset of the patient's psychiatric illness.

EARLIER STUDIES OF GROUPS OF CASES WITH MISSILE WOUNDS OF THE BRAIN

The study of cases of penetrating missile wounds has the advantages that there is usually no preexisting pathology, and that the distribution of lesions is nearly random over both cerebral hemispheres. Most importantly, we can survey relatively large groups of cases that are unselected with regard to symptom. These cases can be studied because they have the penetration into the cranial cavity and not because of any symptoms that might be more or less severe, or more or less noticeable, depending on the lobe and subregion involved.

Feuchtwanger's Work. The classical study of this kind was Feuchtwanger's survey (1923), five years after World War I, of 200 frontal gunshot-

wound cases, systematically contrasted with 200 cases whose known gunshot wounds had entered the skull outside of the frontal region. A careful reading of Feuchtwanger's monograph would have prepared us for the surprises of the forties and fifties of this century: the increasing evidence that frontal lesions, on the whole, have less effect on test intelligence, on categorizing, and on a variety of complex reaction-time tasks than lesions in other areas of the human brain. He concluded that specific changes were not to be sought, either in the intellectual sphere, or in that of "attention" or "memory," including recent memory. Instead, the more pervasive changes were those of mood and attitude, ranging from euphoria, or, less frequently, depression to a curious form of "other-directedness" on the patient's part—an incapacity for making plans. These changes were independent, in Feuchtwanger's view, from the occasional alterations in motor control: restlessness, hyperkinetic behavior, or, more rarely, slowing and torpor.

None of the changes, however, was obligatory. Even Feuchtwanger, with his numerous tests and detailed life histories, found 13 patients among his 200 cases of frontal gunshot wound who were negative in every respect; i.e., they had the lesions, but no symptoms. Similarly, the more characteristic changes in Feuchtwanger's list were actually far from universal: facetiousness, which Feuchtwanger considered an outstanding "frontal" symptom if present, appeared in 9 of 200 frontal cases, and in 1 of the 200 nonfrontals. Such results suggest that in gunshot wounds, with their haphazard and often minimal effects on the injured tissue, positive symptoms may be more revealing than the absence of symp-

toms—a conclusion we have reached somewhat reluctantly over the years.

Interpretations of Earlier Studies of Frontal Trauma. Feuchtwanger's interpretations of his own positive findings are perhaps less helpful. He insisted that psychological changes must be described in the context of an introspectionism in which he had been brought up. Symptoms had to be referred to alterations either in "representations," in "feelings," or in "will." He decided the principal change was in the sphere of voluntary action, and in the ability to evaluate situations (Werturteile).

He insisted that frontal lobe syndromes could not be appropriately described in the framework of purely "associationistic" or "sensationalistic" psychology. If he were with us, and unreformed, he might say that the objective behaviorist can note the changes after frontal lesions but cannot understand them, since he himself has "lost" the problems of "will" and "values"—as being outside the realm of the scientific study of behavior. Yet it is precisely this problem of finding an objective approach to such dimensions of behavior that makes the study of frontal syndromes a particular challenge for experimental psychology.

There is, of course, another way of dealing with frontal symptoms. The way was open to Feuchtwanger, who knew of Kleist's interesting attempt (1931, 1934) to interpret all frontal lobe changes as "psychomotor" symptoms, in line with Wernicke's conception (1874, 1894) of precentral structures (Kleist, 1908). It is curious that Feuchtwanger rejected this approach, in spite of the fact that his results were close to Kleist's descriptions (1931) of "Antriebstörung" (change in drive) as a basic change after fron-

tal lesions, and in spite of the fact that Feuchtwanger had documented a characteristic change in postural control as part of the frontal symptom complex.

The Search for a Theory

SOME CLINICAL OBSERVATIONS ON PROGRESSION OF "FRONTAL SIGNS"

In trying to deal with the challenge raised by Feuchtwanger's results and by our own rather puzzling data, one should turn to some consideration of what is seen in a "typical" clinical case. Take, for instance, the progression of symptoms in a slowly developing bifrontal abscess: The very first changes, often missed, may be indolence, with mild inattentiveness, or dreamy attitudes; later, there may be explosive laughter, or giggles, or, less often, crying spells. There is likely to be increased urinary urgency, and some rigidity of the patient's face, with difficulties in voluntary shifting of gaze. At still later stages, the patient shows the peculiar "Gegenhalten" (cf. Kleist, 1931); he holds on to the examiner's finger; as the examiner moves his own, the patient's rigidity dissolves and he follows smoothly into a new position; wherever the examiner stops, the rigidity immediately reappears. With a few repetitions, the patient gives the impression that he anticipates the end positions, as if he were playing a game with the examiner. This behavior (following, then resisting) is rather different from ordinary Parkinsonian rigidity, which exists over the whole range of motion.

In still more advanced stages of the disease, we have full-fledged, forced grasping, in which the slightest contact, especially that on the inside of the patient's fingers, triggers a com-

pulsive grasp (Janischewsky, 1909; Richter & Hines, 1934). In its severe forms, the grasp reaction is released by vision (Goldstein, 1931), and the patient shows not only forced grasping but groping ("Zwangsgreifen" and "Nachgreifen") following the examiner's hand, or the reflex hammer—even without tactile contact (Rushworth & Denny-Brown, 1959; Schuster & Pinéas, 1926). In discussing these phenomena, Denny-Brown (1951) pointed out that contact responses which precede or accompany a grasping movement have become compulsory stimuli. How can this be interpreted?

STIMULATION OF FRONTAL EYE FIELDS

Or take another simple clinical observation: A patient is on the operating table with bilateral trephine holes over the lateral frontal regions (unpublished observations with F. Echlin). Stimulation on one side, say the left, will induce head turning to the opposite side, but the eyes remain fixed (forward), so that they move relative to their orbits in a compensatory fashion, maintaining fixation during changes of head or body posture. If the head is prevented from moving during such unilateral cortical stimulation, then the eyes move instead to the contralateral side. Similarly, if the room lights are turned off, both head and eyes move to the side opposite the electric stimulus.

This demonstration is, of course, nothing but the classical experiment on the role of "frontal eye fields," reported in detail for waking or lightly anaesthetized monkeys by Graham Brown (1922). He believed the effect indicated that the stimulus activated a highest-level oculomotor mechanism, a "voluntary" level, in contrast to the

automatic or "reflex" fixation mechanism in the preoccipital eye fields (Brown, 1922). Others have attributed the complexity of the frontal-stimulation effect to the presence of a compensatory reaction, while Denny-Brown (1951) prefers an interpretation in terms of general—not purely visual—orienting responses. But are these interpretations incompatible? I rather suspect that a great many of the frontal symptoms can be understood as falling into the same general class of defects in compensatory mechanisms which normally permit highest-level motor functions in response to sensory stimulation. Thus, the old-fashioned view we have gradually adopted is that frontal lobe lesions produce changes on the sensorimotor level, but changes of a very special sort.

SKETCH FOR A THEORY: ROLE OF FRONTAL STRUCTURES IN MAINTAINING A "COROLLARY DISCHARGE"

What is required here is the 180° turn in our approach, the turn we have promised from the beginning. Perhaps the greatest difficulty in understanding the frontal lobes has been the classical tendency to start all considerations of brain function from the sensory side, and to proceed in the familiar fashion from the sensory to the motor. We propose instead that an important contribution of frontal structures is to the corollary discharge—i.e., to a discharge from motor to sensory structures—which prepares the sensory structures for an anticipated change (Teuber, 1960, 1961a, 1961b). Thus, a grasp will lead to an unexpected contact sensation which signals the successful completion of the act of grasping by virtue of its matching the prediction of contact,

which thus becomes both a trigger and a brake of the normal motor act. The principle can be made even clearer in considering the control of voluntary eye movement (Teuber, 1960, 1961a, 1961b; von Holst & Mittelstaedt, 1950). During a voluntary shift of gaze, the environment stands still—presumably because a corollary discharge from the oculomotor to the visual mechanism prepares the latter for the change in relative position which will result from the ocular motion (Sperry, 1950). Passive motion of the eyeball, as can be produced by pushing against the eye, makes the world move, because the counterbalancing corollary discharge is missing. Similarly, inverting spectacles (Held, 1961; Kohler, 1951) lead to exaggerated motions of the visual field, with every movement of the eyes or head, because the corollary discharge has the wrong sign. Most importantly, paralysis of eye movement (as in external rectus palsy) is experienced by the patient as a continuous moving about of the visual scene, because the discharge from motor to sensory systems is now running off *in vacuo*, unchecked by the normal motor effect.

We have come to the conclusion that disturbances on this level of motor function, i.e., on the level of the corollary or anticipatory discharge, represent an important common denominator in many forms of frontal lobe pathology in man. This recognition has arisen against the background of an unexpectedly large body of negative evidence, i.e., of tasks on which our patients with frontal lesions either performed normally or at least no worse than patients with lesions elsewhere in the brain. Both the negative evidence and the few positive symptoms to which we shall turn later have to be considered within the framework

of the particular patient population in which we undertook our search for specific and lasting changes after frontal lesions.

Experimental Results

GROUPS STUDIED BY US

As in our work on occipital, parietal and temporal syndromes, the effects of frontal lesions were observed in the course of studying 232 men with gunshot wounds of the brain, and 118 controls with injury of peripheral nerves, but not of the brain. The selection of these cases, their grouping into classes determined by presumed site of lesion (locus classification) or by various quantified sensory or motor deficits (symptom classification) have been repeatedly described (Teuber, 1959; Teuber et al., 1960). Crucial to our selection of cases was the requirement that the men be seen by us because they had the lesions and not because they did or did not have any particular need for clinical attention. Of the 232 men, 90 had frontal penetration (42 left, 22 right, and 26 bilateral). Clinically, only very few men in this subgroup presented gross frontal lobe syndromes, such as are seen with space-occupying lesions of the frontal structures.

NEGATIVE EVIDENCE: GENERAL OBSERVATIONS

There is thus far no indication that the frontal group compares unfavorably with the other locus groupings in terms of post-injury employment, or domestic arrangements. Only very few of the frontal changes involve difficulties of the sort usually cited in the literature. Thus, one man has developed a peculiar mythomania—telling of various imaginary exploits,

such as hurried trips to the moon between testing sessions. In this case, the pathological lying has led to dismissal from his job as bank clerk, although he did not to our knowledge show other forms of unconventional behavior. One doubts, of course, whether this single case among so many justifies attribution of the behavioral disturbance to the injury, but the telling of tall tales, often without any obvious striving for gain—has occasionally been described as a sign of frontal lesions.

Two men show marked facetiousness; interestingly, one of them is also hyperkinetic, constantly pacing about the room, talking loudly, with explosive laughter; when he sits down, he continuously picks imaginary specks from his clothes or from the furniture. The other, however, shows facetiousness against a background of diminished motility, with an appearance of mild depression. Neither of these two is really employable. On the other hand, there are numerous patients with excellent post-injury adjustment, including one lawyer, one certified accountant (who acquired his training after a severe bifrontal injury), and a minister who also trained only after the injury. Thus, the same puzzling rarity and randomness of severe change is found in this group as one would find in others, provided the groups are not selected for the presence of "trouble."

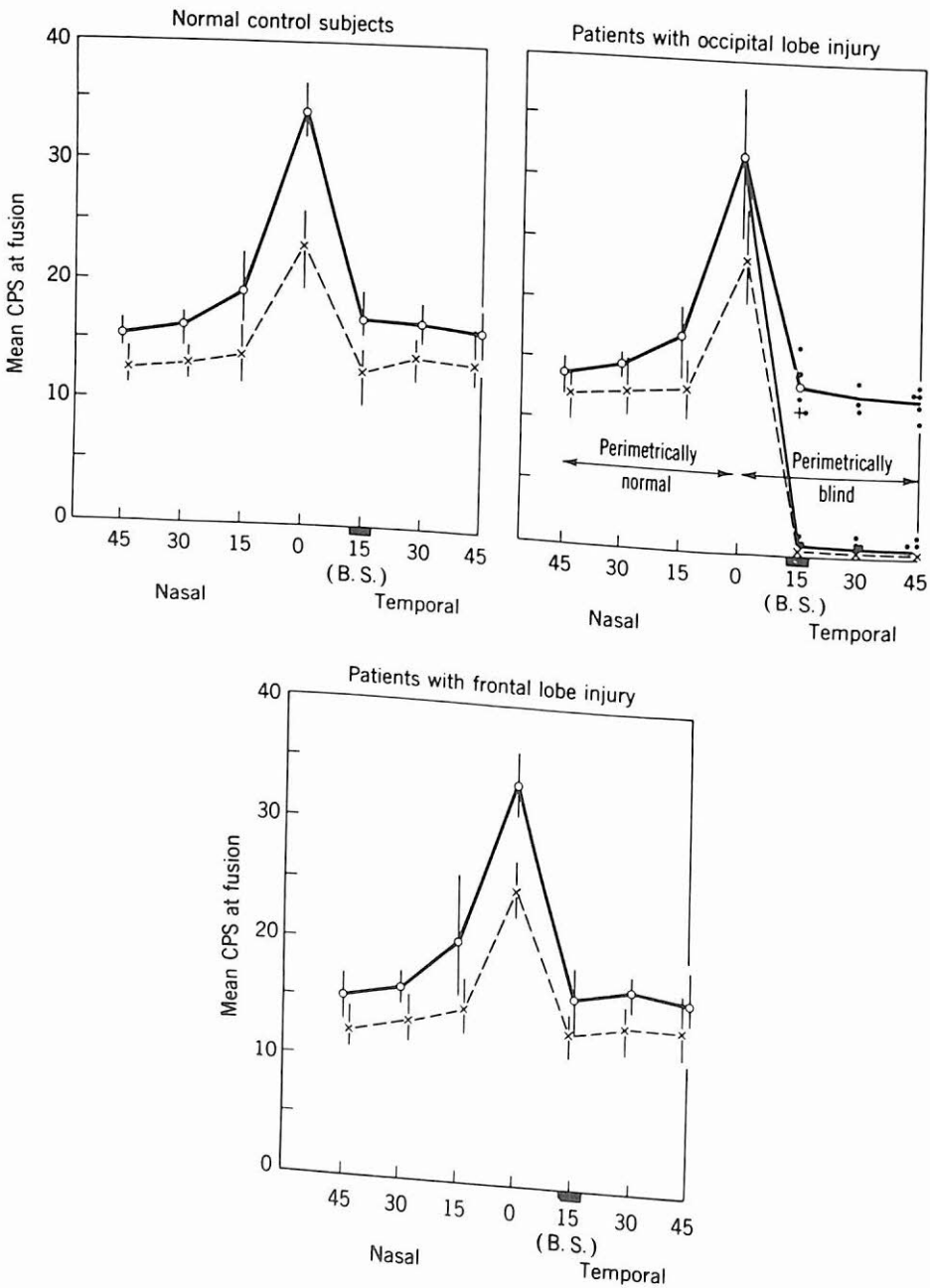
NEGATIVE EVIDENCE: SPECIFIC EXPERIMENTAL RESULTS

CFF. A similar preponderance of negative results appears on formal testing. Fusion thresholds for flicker-light (CFF) have been studied extensively in this group, particularly by Battersby (1951; Battersby et al., 1951; Teuber & Bender, 1948; Teuber

et al., 1960). Regardless of the method employed, whether with large or small targets, in the center or periphery of the field, with monocular or binocular regard, the frontal group shows normal or slightly better than normal fusion thresholds. By contrast, our occipital lobe group, as expected, shows a significant and permanent reduction in CFF (Figure 20.2).

Sorting Tests and Related Techniques. In our hands, tasks such as the modified Weigl Card Sorting Test (Weigl, 1941) in the Wisconsin version (Grant & Berg, 1948) did reveal a deficit in frontals, as compared with controls, but a slightly more severe deficit appeared after parieto-temporal lesions—especially on the left (Teuber et al., 1951). The deficit is thus non-specific for frontal lesions—at least with the method of presentation we used. The same can be said for the numerous variations of these tasks employed in our laboratory, including essentially nonverbal versions of oddity, matching-from-sample, and conditional matching tasks (Teuber et al., 1951; Semmes et al., 1954; Weinstein et al., 1955). All three tests were administered in two versions: through the visual and through the tactile modality. As we have reported elsewhere (Semmes et al., 1954; Weinstein et al., 1955), maximal defects appear either with parietal or left parieto-temporal lesions, suggesting again that tasks of this type are not likely to uncover specifically frontal deficits against the background of essentially unselected cases with lesions in any part of the brain.

AGCT. The same must be said for results on tests of general intelligence, such as the Army General Classification Test (AGCT). Here too, maximal impairment appeared after left parietotemporal lesions (Teuber,



Critical flicker frequency (CFF) as determined by flicker perimetry for 10 normal control subjects (left), 10 patients with lesions of occipital lobe and resulting homonymous hemianopia (right graph), and 10 men with frontal lobe injury

Figure 20.2

1959; Teuber & Weinstein, 1958; Weinstein & Teuber, 1957), rather than frontal—a result sharply at variance with Rylander's reports (1939, 1943), but in line with results for tumor cases obtained more recently by Pollack (1955) and others.

Problem-solving Behavior. Tests of problem solving in somewhat more lifelike situations were equally unproductive of specifically "frontal" deficits. The tasks employed (Battersby et al., 1953) were patterned after those suggested by Maier (1931, 1945) for the investigation of practical reasoning. Here again, patients with frontal penetration did not do any worse than those with lesions in the posterior brain substance.

Formboard. In contrast to earlier reports by Halstead (1947) was the finding that formboard and formboard recall scores were also lower after posterior lobe, rather than frontal lesions (Figure 20.3; Teuber & Weinstein, 1954). This again is not unreasonable, in view of the known symptomatology of parietal and temporal lesions. What is difficult to un-

derstand is that this formboard task should have been considered a test of frontal pathology at all.

Special Tests of Recent Memory. To continue with such negative results: Drs. Ghent, Mishkin, and I (1962) have recently completed a study representing a concerted attempt at teasing out defects in recent memory in our frontal group. Various tests of recent memory were devised for this study, in which we tried to provide tasks not readily memorized by verbal means. For instance, patients were required to recall, with or without a 15-sec delay, the tilt of a line that had been exposed in the dark; the position of small points of light, in different parts of their visual field; the position of punctate contacts on their own palm; or, lastly, their own body position after brief periods of tilt in a tilting chair. On all these tasks, the frontals failed to show any deficits, as compared with controls. Such a negative finding gains in importance if one considers that the posterior-lesion groups did show abnormalities on some of the tasks. Conventional tests

(lower graph). The average cycles per sec at fusion are given on the ordinate. The positions of the flickering patch in the subject's field are given along the abscissa, with O indicating the central fovea, and 15, 30, and 45, respectively, the distances (in degrees of arc) on the midhorizontal meridian, in the nasal or temporal half of the field. B.S. refers to the region of the normal blind spot; note that CFF in that region is only slightly reduced, being mediated by scattered light around the optic disc. The targets employed were 2° in diameter (solid line in graph) and $\frac{1}{2}^\circ$ (interrupted line); the light source (a Sylvania tube R 1130B) emitted flashes at a light-dark ratio of 1, and a peak intensity of 8.92 a.f.c. ft². The lengths of vertical bars indicate scatter of thresholds obtained (one standard deviation above, and one below the mean). Note that half of the 10 men with occipital lesions (center group) showed some residual light perception in that half of their field which seemed totally blind on perimetric testing. By contrast, the (perimetrically) "normal" half of the field in the occipital lesion cases showed a small but significant reduction in CFF. Note also that CFF values for the frontal lesion group (lower graph) are essentially normal. (Graphs based on Battersby, 1951, and Teuber et al., 1960.)

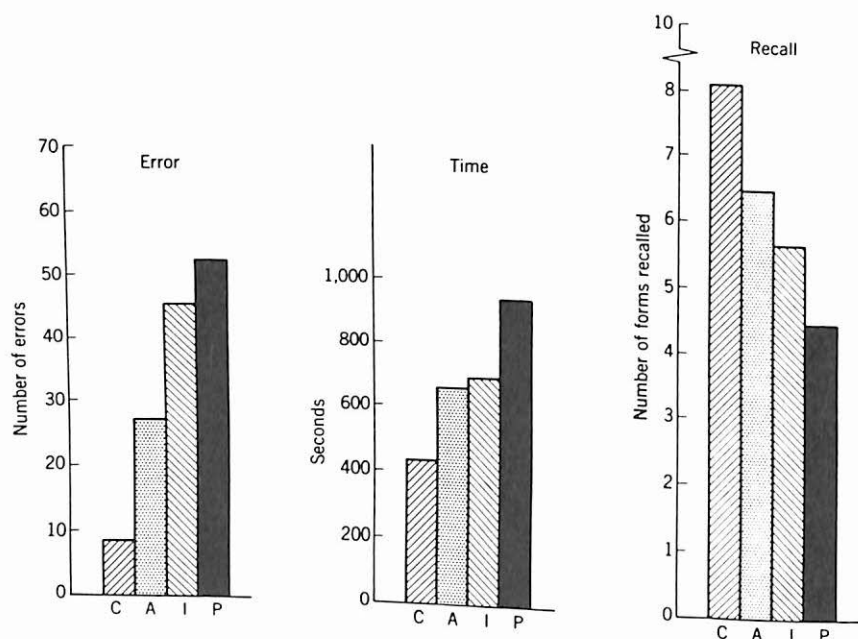


Figure 20.3

Results of formboard test (Seguin-Goddard formboard) in 12 normal controls (C), 17 men with brain wounds implicating the anterior third of the brain (A), 15 with wounds of the intermediate third (I), and 15 with wounds of the posterior third (P). The blindfolded subject attempts to fit 10 different forms into the corresponding 10 holes in the board, first with the board in its usual position, then with the board inverted (rotated 180°). The total error scores (all "false entries," i.e., attempts at fitting a form into an incorrect opening in the board) refer to a sum for both positions of the board, as do the scores for time (in sec) and recall (number of forms correctly drawn from memory after the main task has been completed). On all three scores (errors, time, recall) brain-injured subgroups are significantly inferior to controls; on recall, the group with anterior lesions did significantly better than the other brain-injured subgroups, and the group with posterior lesions significantly worse. (Based on Teuber & Weinstein, 1954.)

of recent memory (digit span, and a visual variant for this, viz., memory span for visually presented sets of stereometric patterns) were likewise unproductive of any frontal deficit.

Taken together, all our negative results may mean that larger lesions, or acute lesions, are needed to uncover major frontal lobe pathology in man;

yet this conclusion is made unlikely by the fact that a few rather special tasks did bring out lasting symptoms after gunshot wound of the frontal lobes. None of these symptoms, however, can be described as primarily intellectual, or as defects of attention or memory, no matter how defined. Let us turn, therefore, to those few posi-

tive findings that have accumulated in our work over the last 12 years.

POSITIVE FINDINGS

Visual Searching Tests. To begin with, we have definite evidence for a subtle but lasting deficit in visual searching (Teuber et al., 1949). The test employed is an elaboration of a technique originally used by Poppelreuter (1917)—the “Field of Search” (Figure 20.4). The task involves presentation of an array of 48 patterns distributed irregularly over a screen

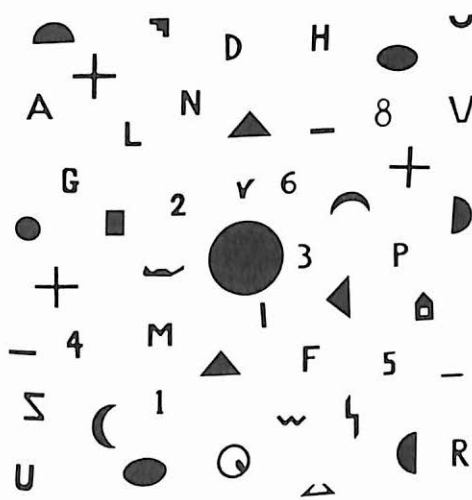


Figure 20.4

One of six test fields used in “field of search” procedures. Forty-eight patterns (distinguished by either shape or color, or both) are shown on a large rear-projection screen, so that the array of patterns subtends 60° by 60° of the subject’s field. The subject is seated in front of the display forming the circular center area (here shown in black). At a warning signal a duplicate of one of the 48 patterns appears in the center, and the subject has to find the matching pattern in the periphery, with active eye and head movements, and identify it by pointing.

which subtends 60° by 60° of the patient’s visual field. In the center of the field, a separate projector can show a duplicate of any one of the 48 patterns, after a warning signal. Immediately after the appearance of the duplicate test pattern, the patient searches with active eye and head movement for that pattern in the periphery that matches the one in the center. He is timed for each of the search periods, and the results are plotted separately for each quadrant of the field. The effect of frontal lesions is a significant reduction in speed and efficiency of search—hence prolonged searching times, and, in unilateral cases, a marked asymmetry of searching times, so that the search for objects opposite the lesion is disproportionately prolonged (Figure 20.5). These asymmetries persist in the absence of any clinically obvious paresis of lateral gaze. In fact, there are indications that the symptom is maximal whenever the patient tries to fixate and is less pronounced when the search occurs with freely moving gaze. It seems as if the symptom reflected an overactivity of the posterior (preoccipital) fixation mechanism, in the presence of damage to areas in and around the so-called frontal eye fields (see also, Bianchi, 1895; Kennard, 1939; Kennard & Ectors, 1938; Loeb, 1884; Silberpfennig, 1941; Welch & Stuteville, 1958).

It must be added, however, that some disturbance of searching can be found with posterior-parietal and with occipital lesions. We are hoping to show by recording the actual searching movements in either case that the posterior deficit is different in kind, being primarily a derangement of fixation as such (Denny-Brown, 1951).

The Aubert Task and Its Variants. A curious aftereffect of anterior-lobe

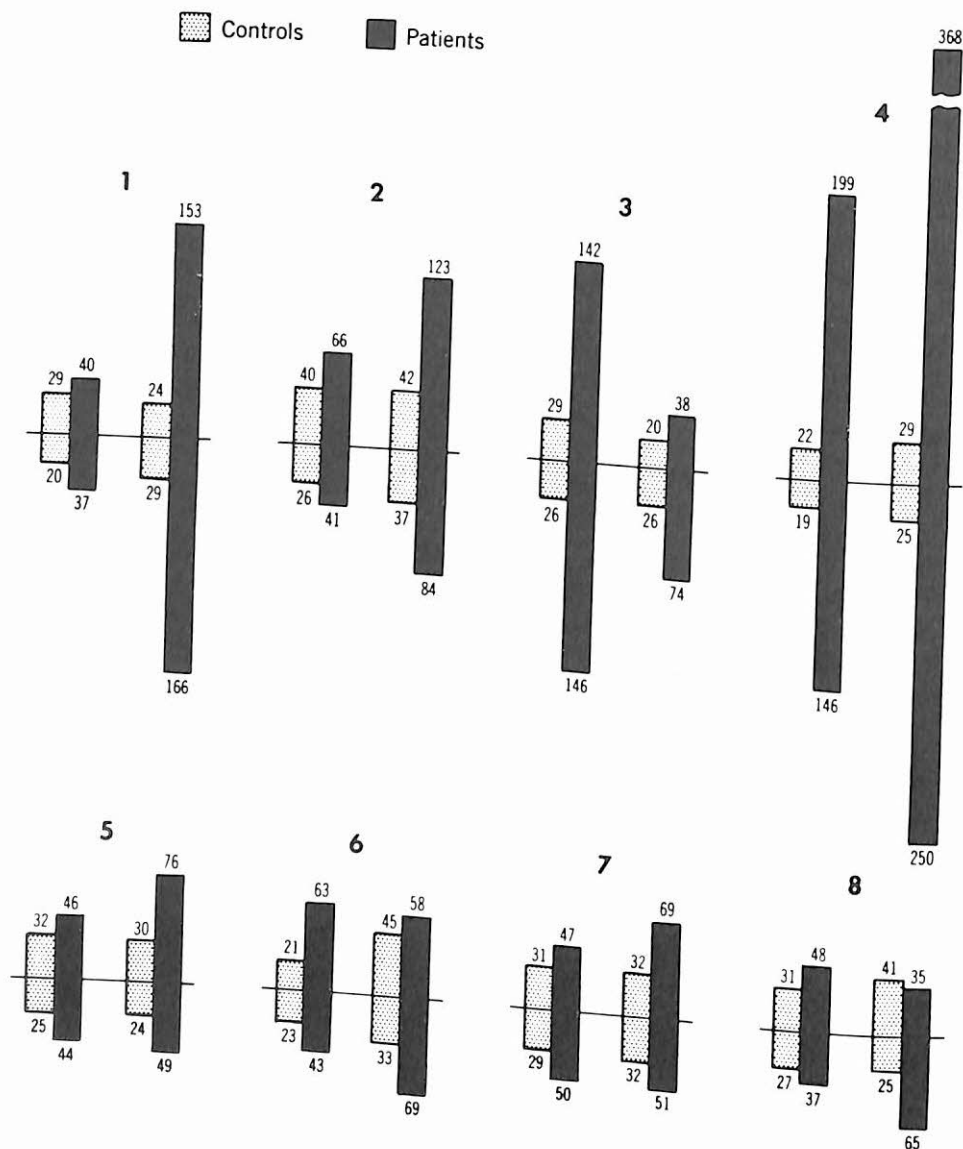


Figure 20.5
Illustrative results of "field of search" tests in 16 cases (8 controls and 8 brain-injured). The bar diagrams indicate total searching times (in secs) for each quadrant of the field—upper and lower left, and upper and lower right, respectively. Cases 1 and 2 had lesion of the left frontal lobe, case 3, right frontal, and case 4, bilateral frontal penetration. Cases 5 through 8 had penetrating wounds of rolandic, temporoparietal, parietal, and occipital regions, respectively. Note prolongation of searching times in the cases of primarily frontal injury, particularly in the half-fields opposite to their lesions. (Based on Teuber & Battersby, 1949, unpublished observations.)

penetration in man is the perceptual dysfunction described by Mort Mishkin and me, nearly 10 years ago (Teuber & Mishkin, 1954). The phenomenon itself is simple. When I tilt my head in a lighted room, the vertical lines between floor and ceiling do not tilt—they remain upright. The mechanism which makes this possible works in the dark, but not as well. Tilting one's head, or head and body, in the dark while looking at a luminous line (Aubert, 1861), leads to errors in attempts at setting the line to the vertical. With moderate body tilts (up to 30°) the error consists in bringing the test line too far opposite to one's tilt; the line will be inclined away from true vertical by a few degrees to the left with body tilts to the right, and conversely with body tilts to the left.

We undertook to analyze this phenomenon by a series of experiments on the chance that frontal lesions might affect it. First we let patients and controls set a luminous line to the vertical while they sat upright in a chair (Figure 20.6, experiment 1). Then the same judgments were required of them with body and head tilted 30° to the left, and after that, 30° to the right (Figure 20.6, experiment 2). The subjects were 20 of our frontal patients, 20 of those with parietal lesions, and 20 controls with peripheral nerve injuries (Figure 20.7). The results were clear: As long as head and body were upright (experiment 1) there were only very slight errors in any of the three groups, and these errors did not differentiate the patients from the controls. With body tilts to the left or right, however, the anterior lobe patients made exaggerated compensatory errors (Figure 20.8), and these errors were larger than those of the posterior-lesion cases, who in turn made larger

errors than the controls. We interpreted this result (Teuber & Mishkin, 1954) as an exaggeration in the effect of those compensatory mechanisms that provide normal distance perception under conditions of changing posture—a reaction to visuopostural conflict.

Conflict of this sort but within a single modality was provided by the next experiment (Figure 20.6, experiment 3). Here the patient was required to set a line to the vertical against a conflicting visual background, without body tilt. The results were, in a way, the reverse of those for experiment 2, i.e., for the visuopostural conflict situation (Figure 20.9). Now the posterior-lesion group did worse than the frontals, who came closer to the controls. This visuovisual conflict did not suffice to bring out a frontal deficit.

Finally, we asked whether purely postural factors, in the absence of distance-receptor action, might account for our findings. Thus, we did experiment 4 (Figure 20.6, experiment 4). Here the patient was merely required to bring himself back to the postural vertical while blindfolded, after varying periods of tilt on the tilting chair. This fourth experiment disclosed a moderate constant-error tendency in all groups: Patients as well as controls tended to stop short of true vertical to the side of a previously maintained body tilt. Thus, this last (fourth) experiment—just as the first of the series—was nondifferential in its results, indicating that neither purely visual, nor purely postural factors could account for the abnormal reaction of the frontals on the visuopostural task.

All four experiments can be schematized in a single figure, permitting us to summarize the results (Figure 20.6). The first and last experiments—visual per se and postural per se—have no

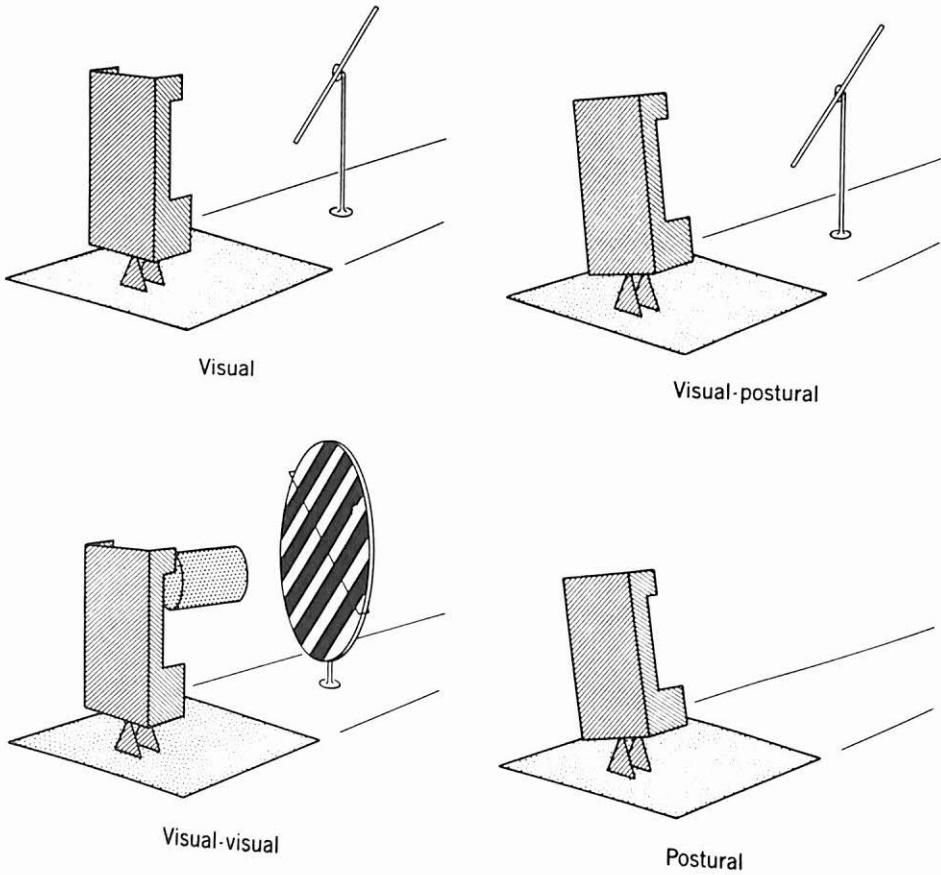


Figure 20.6

Diagrams showing sequence of experiments involving judgments of visual or postural vertical, with body upright (experiments 1 and 3), or body tilted (experiments 2 and 4). In experiment 1, the subject is upright and sets a luminous line in a dark room to apparent vertical, under purely visual control. In experiment 2, the same visual judgment is made under conditions of abnormal posture (body tilted), hence "visuopostural conflict." In experiment 3, posture is normal, but a conflicting visual background (oblique stripes) complicates the subject's attempts at setting a black thread to the apparent vertical. In experiment 4, the subject is blindfolded and has to bring himself back to the apparent postural vertical, hence a purely postural task. (After Teuber & Mishkin, 1954.)

differential outcome. It should be noted, however, that in the acute phase after frontal lesions, there may be errors in the setting of the vertical with body upright (Bender & Jung, 1948), and there may also be transient abnor-

malities of posture which disappear in a matter of weeks or months (Bruns, 1892; Delmas-Marsalat, 1936; Feuchtwanger, 1923; Gerstmann, 1916). The second experiment (visuopostural interaction in setting the line with body

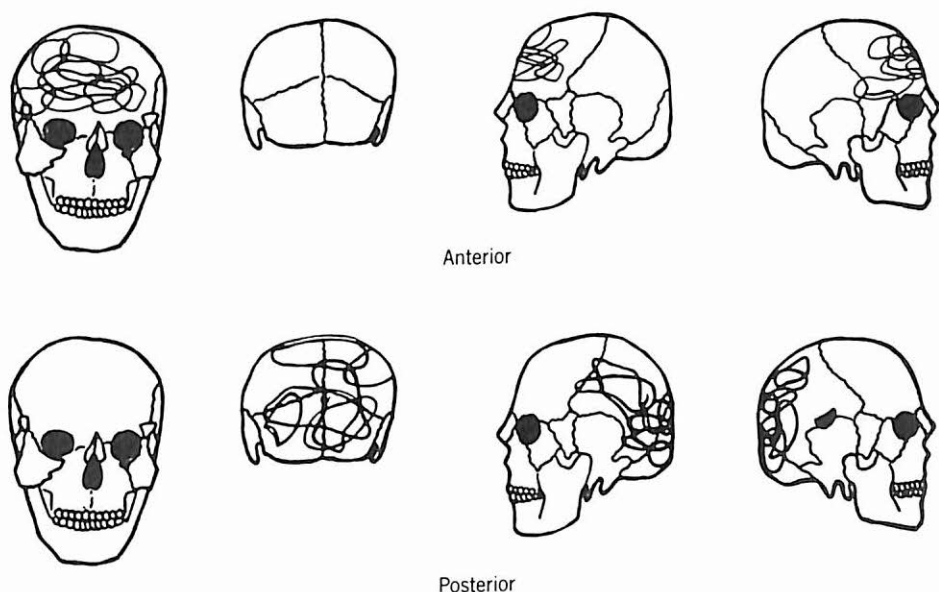


Figure 20.7

Composite diagrams showing outlines of skull defects in 20 men with anterior (frontal) and 20 with posterior (parieto-occipital) missile wounds of brain. These 40 men participated in the four experiments shown schematically in Figure 20.6. (From Teuber & Mishkin, 1954.)

tilted) yields a primarily frontal sign, and the third experiment (here labeled visuovisual, setting a thread to the vertical against a conflicting, obliquely striped background) yields a posterior parietal sign. The opposite results of these experiments provide an especially striking case of what we have called "double dissociation" of symptoms (Teuber, 1955).

Incidentally, normal children do not show the compensatory errors made by adults on the tilting chair, and made in such exaggerated fashion by the adults with frontal lobe lesions. Nor does early brain injury in children produce the exaggerated error tendency of the adult frontal lobe syndrome. Instead, brain-damaged children differ from normal children matched for age

by showing less error, i.e., by being paradoxically better on this task (Teuber & Rudel, 1962).

A further point about the tilting chair experiments. They have been extended quite recently by Florry Proctor (1961), in our laboratory, to patients with basal ganglia pathology (Parkinsonism) on Dr. Irving Cooper's service at St. Barnabas Hospital in New York. These patients made the same exaggerated constant errors on setting the visual vertical while their body was tilted; they resembled, in this respect, our cases of frontal gunshot wound in young adults. In addition, however, the Parkinsonian cases also made markedly exaggerated errors on the purely postural task, the fourth experiment of the series, on

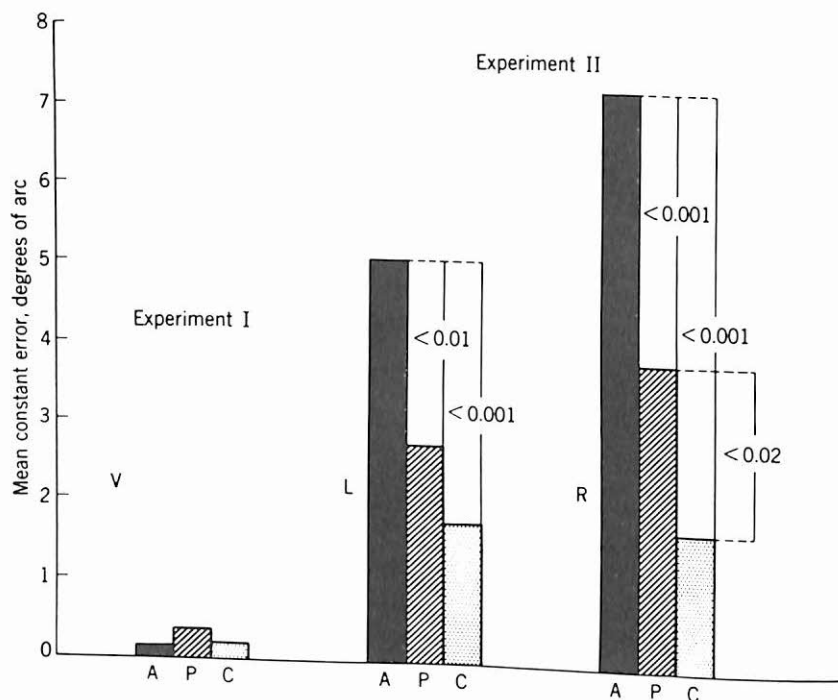


Figure 20.8 Results (average constant errors of tasks involving the setting of the visual vertical with normal posture marked V (body vertical, experiment 1) and with abnormal posture (experiment 2, i.e., body tilted 28° to the subject's left—diagrams marked L, and 28° to the subject's right—diagrams marked R). Labels A, P, and C, beneath the bar diagrams, refer to 20 controls (C), 20 men with anterior lesions (A), and 20 with posterior lesions (P), as shown in Figure 20.7. Significant differences between experimental groups are shown next to the appropriate parts of the histograms. For further details, see text. (From Teuber & Mishkin, 1954.)

which our traumatic frontal cases had not shown any abnormality. Thus, the sequence of experiments revealed both a similarity and a difference in adults between frontal trauma and basal ganglia disease.

Lastly, we should mention that the results obtained by Mishkin and me have been subjected to yet another analysis. Instead of extracting the constant (average) errors of setting, one can compute the so-called "starting-position" errors; in this way one can disclose a completely nonspecific effect

of brain injury (Teuber & Liebert, 1958). These starting-position errors are those related to the starting position of the line in any given trial. Thus, if the luminous line is brought up toward the vertical from the patient's left, most patients tend to set it somewhat farther to the left than on those trials where the line is brought up toward the vertical from the patient's right. On these latter trials the line is set farther to the right. By averaging all trials with alternating starting positions from either side, one

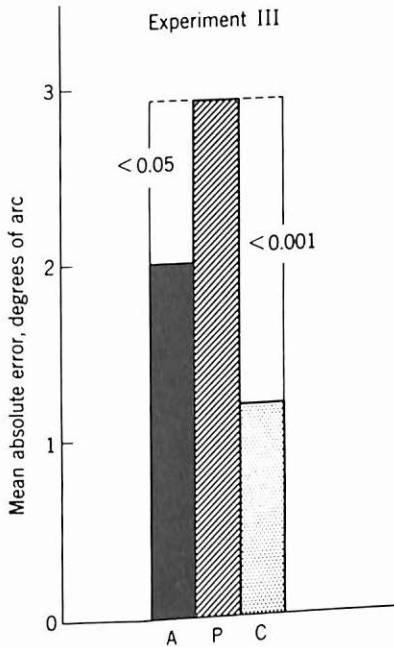


Figure 20.9

Results of settings of visual vertical in experiment 3 (compare Figure 20.6, visual-visual). The subjects attempt to set a black thread to the apparent vertical against a conflicting (obliquely striped) visual background. For further details see text, and legend to Figure 20.8.

obliterates this particular starting-position effect. By contrast, assessing the effects separately—for starting positions from the right and for those from the left—will yield two error scores which are greater in all brain-injured subgroups than in the controls, yet fail to distinguish among brain-injured subgroups.

We thus derive from one series of experiments both specific (localizable) and general symptoms, depending merely on the way in which the results are analyzed. Starting-position effects in this and similar situations are correlated with performance on hidden-

figure tasks (Teuber & Weinstein, 1956), which are also done poorly by our brain-injured patients regardless of site of lesion—whether the lesions involve the frontal, parietal, temporal, or occipital lobes. These starting-position effects are also rather large in normal children, and still larger in brain-injured children of any age (Teuber & Rudel, 1962).

Reversible Figures. In a paper on the use of hidden figures with children, Lila Ghent (1956) suggested that reversible figures such as the Necker cube or similar patterns with ambiguous perspective (Figure 20.10) might be considered an extreme case of hidden figures. Instead of embedding a figure among the concealing contours of another, an identical set of contours is now seen as pertaining to one, and now to another perspective pattern. However, the perception of sudden reversals in such patterns harbors another element: that of a virtual movement into depth, now this way, now that, and a correspondingly implicit shift in the observer's standpoint relative to the three-dimensional pattern perceived. These considerations, to-

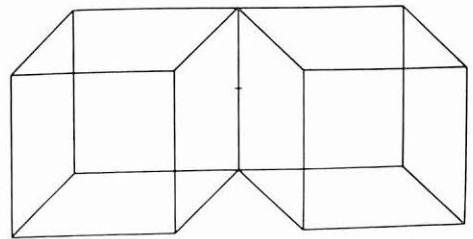


Figure 20.10

Double Necker cube, employed in studying reversals of ambiguous perspective (see Cohen, 1959, and Teuber, 1960). The subject fixates the intersection point between the short horizontal bar and the vertical line bisecting the total figure.

gether with earlier work by others (Gelb & Goldstein, 1920; Harrower, 1939; Yacorzynski et al., 1948; Petrie, 1952a) made us suspect that perspective reversals might be abnormal after brain lesions and maximally so with frontal lobe involvement.

The systematic experiments of Leon Cohen in our laboratory (1959) have confirmed this suspicion but have gone beyond it by demonstrating an unexpected difference between effects of unilateral and bilateral frontal gunshot wounds. He used for his main series of tests a double Necker cube (Figure 20.10). The patient had a tapping board with two keys in front of him, and struck one key, to the left, when the left cube seemed to reverse itself; he struck the right key for a reversal of the cube on the right, and both keys when the cubes seemed to reverse themselves together. A given experimental session lasted 90 sec and the cumulative reversals were graphed for the six consecutive 15-sec periods that made up the minute and a half of continuous observation.

The normal controls showed a characteristic pattern: first, rapid acceleration of reversals, then a tapering off (Figure 20.11). The men with unilateral lesions in lobes other than frontal showed fewer reversals, and those with biparietal lesions showed particularly few. Yet the unilateral frontal-lesion groups (right or left) differed significantly from the controls and all other brain-injured by showing the fewest reversals (Figure 20.11). The most curious aspect of these results, however, was the performance of the bilateral frontals. They too differed from the controls, and significantly so, but they differed by going in the opposite direction. The bifrontal group showed more rapid reversals than the controls or any of the other brain-

injured subgroups (Figure 20.11). This last result was rather unexpected and remains unexplained. Perhaps it means that partial (unilateral) lesions promote a compensatory tendency to "hold" one perspective, and hence to slow down reversals, while bilateral lesions in the frontal lobe substance induce the abnormally rapid oscillation.

If we had averaged all the results for the frontal group, the opposite changes for unilateral and bilateral lesions would have canceled each other, and only the greater range of scores would have distinguished frontals from nonfrontals. For this reason, we are now reanalyzing other results of tests given to our frontal and non-frontal groups to see whether separate treatment of unilateral and bilateral lesions might disclose differences not found in earlier computations.

It is, of course, possible that our bilateral frontal lesions are not equivalent (anatomically) to a combination of the usual unilateral gunshot wounds involving the frontal lobe substance. Thus, the bilateral lesions may implicate less of the orbital parts of the lobe, or more of the medial or deeper structures not ordinarily involved in the unilateral penetrations.

"Personal" Orientation Test. Similar considerations apply to the fourth and, so far, last of the tests which in our hands have distinguished frontal from nonfrontal brain injury. These experiments are about to be published in *Brain* by Semmes et al. (1963). The work is an extension of earlier published findings involving spatial orientation and its changes after cerebral gunshot wound (Semmes et al., 1955). A task such as the one shown in the next figure (Figure 20.12) reveals a specific "parietal" symptom (Figure 20.13). It was unexpected that a seem-

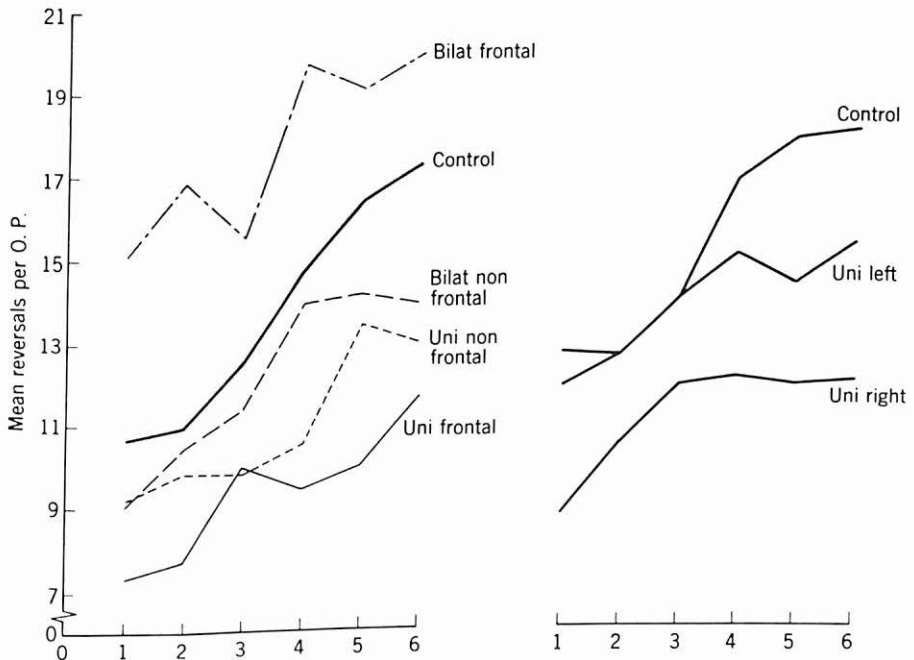


Figure 20.11

Average number of apparent reversals of Necker cubes (shown in preceding figure) over a 90-sec period of continued inspection, grouped into six observation periods of 15 sec each (O.P., along abscissa). Graphs on left side: 19 unilateral frontals (lowest reversal rates); 20 unilateral nonfrontal lesions (parietal, temporal, or occipital); 4 bilateral nonfrontals (occipitoparietal); 21 controls (without brain injury); and 10 men with bilateral frontal lesions (highest reversal rates). On the right side of the graph results of a replication of the experiment are shown for the same 21 controls and for 21 cases of unilateral involvement of the left hemisphere, and for 18 cases of unilateral involvement of the right hemisphere. The right hemisphere group showed markedly fewer reversals than those with injuries in the left cerebral hemisphere. (From Teuber, 1960, based on Cohen, 1959.)

ingly similar task involving orientation to the patient's own body (Figure 20.14) should disclose, at least in part, a "frontal sign."

To turn first to the earlier "spatial orientation" task: here a patient carries a series of maps, one at a time, in a room where nine dots are painted on the floor in an evenly spaced pattern. His task is to pace out on the floor that route which is laid out on a

given map. Performance on this task is clearly impaired with parietal lesions, right, left, or bilateral, as we have described repeatedly elsewhere (Figure 20.13). The poor performance on this route-finding test is correlated with two-point discrimination deficits on one or both hands (Weinstein et al., 1956), again essentially in line with what might be expected, although it must be stressed that the route-finding

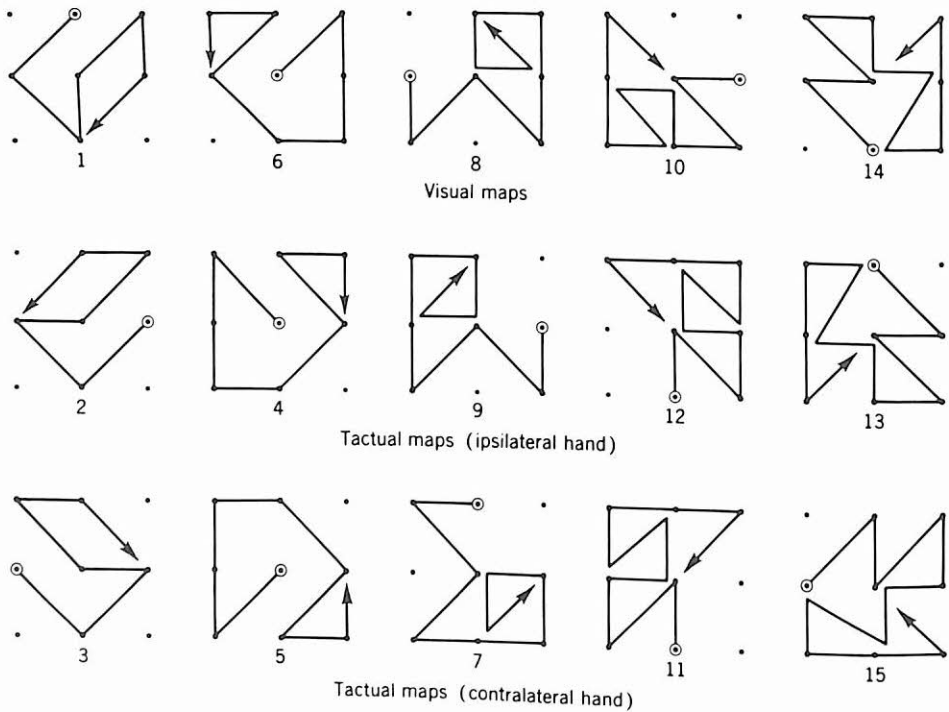


Figure 20.12

Maps used in tests of spatial orientation (extrapersonal orientation) for brain-injured patients and controls. Five maps were presented visually, five tactual, to the hand on the same side as the brain injury, and five to the hand opposite the injury. In the control group, right and left hand were alternately designated as contralateral. The path followed for any given map was scored as correct or incorrect as a whole, but the scores were multiplied for each map according to the number of turns required. In each of the three series of maps (visual, tactual ipsilateral, and tactual contralateral), the number of turns required increased by one from five (first map in series) to nine (fifth map in series). The highest possible score for each series was thus 35. For further details, see text. (From Semmes et al., 1955.)

deficit is independent of the sense modality through which the maps are presented (visual, or tactual, as shown in Figure 20.12). We thought that a task requiring orientation to the body should yield analogous results—that is, maximal deficit with parietal lesions, but this was not so. Figure 20.14 shows the kind of task employed in these new

tests of “personal” (instead of “extrapersonal”) orientation. The patient, standing upright, faced a series of schematic drawings of the human figure, on which one figure represented a man facing the patient, the other a man facing away from him. Numbers appeared on different body parts in the manner shown (Figure 20.14). The patient’s task was to touch, on his own

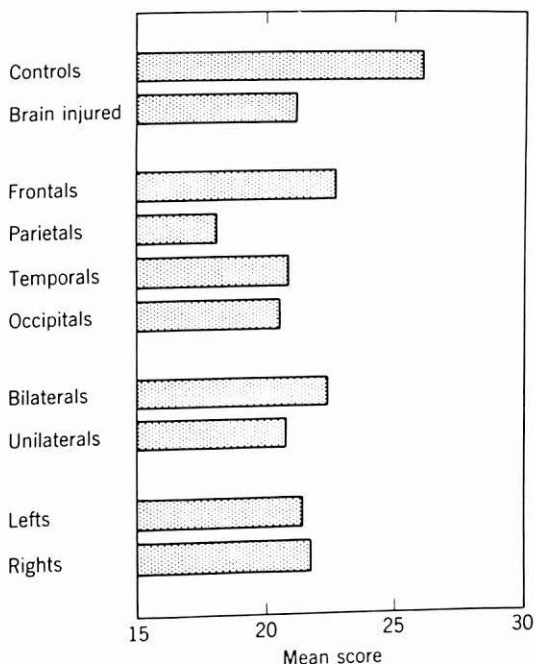


Figure 20.13

Results of tests of spatial orientation (extrapersonal orientation, see preceding figure), irrespective of mode of presentation (visual, tactual ipsilateral, or tactual contralateral), for the groups indicated, with 17 controls, 62 brain-injured men (viz., 14 bilateral lesions, 21 with left, and 27 with right unilateral involvement). The frontal lobe was implicated in 32 of these cases (and spared in 30), the parietal lobe in 23 (spared in 39), the temporal lobe in 19 (spared in 43), and the occipital lobe in 19 (spared in 43). Note that parietal lesions produced differentially greater impairment on these tasks than lesions elsewhere in the brain.

body, those parts that were identical with the parts numbered on the drawings, and he was to do so in the order of the numbers.

The task is clearly an extension of Henry Head's well-known eye-and-ear tests (1926), requiring rapid and repeated shifting from front to back view and the accurate matching of depicted body parts with parts of one's own body. Traditionally, right-left confusion and difficulties with naming or finding body parts are assigned to

trouble in the left parietal region (Ajuriaguerra & Hécaen, 1960; Conrad, 1932; Critchley, 1953; Gerstmann, 1924, 1930; Klein, 1933; Lhermitte, 1939; Nielsen, 1938; Zangwill, 1951). We did find, in fact, that the left posterior group did poorly on this task (Figure 20.15), but so did the left and the right anterior (i.e., frontal) lesion groups. For the right hemisphere but not for the left, we thus obtained double dissociation of personal and extrapersonal orientation—orientation

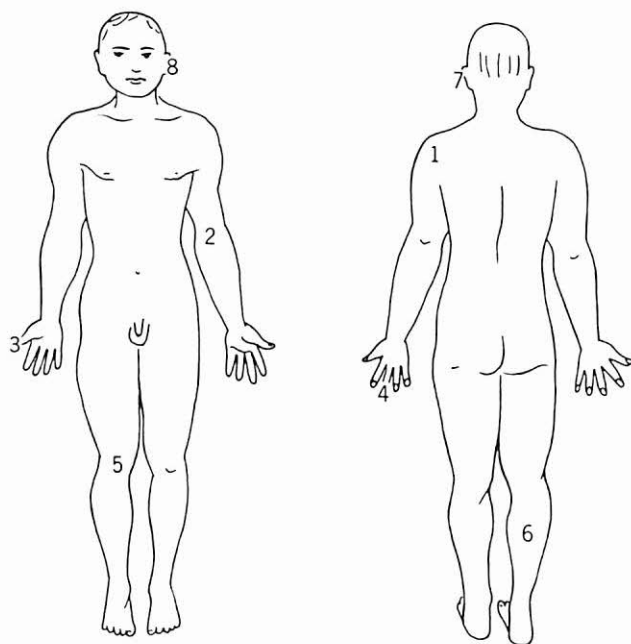


Figure 20.14

Sample figure (one of five) from test of personal orientation (Semmes et al., 1963). Each of the five diagrams in this test consists of a front and back view of a man, with a series of numbers on various body parts, as shown. The subject is required to touch parts of his own body in the order given by the numbers on the diagram.

to surrounding (extrapersonal) space was more vulnerable to posterior lesions, and orientation to the body was more vulnerable to frontal. What is not clear is whether the repeated reversal (of a principle) per se produced the difficulty on this task for our frontals, or whether their main difficulty was more specifically related to orientation and reorientation to the body.

If we adopt the latter interpretation, then the four different frontal symptoms we have found so far—against the background of numerous negative findings in the same group of patients—would seem to fit together and so yield perhaps to a unitary interpretation. The essential alteration would be in some mechanism which

enters into proper direction of voluntary gaze, on searching tasks, and into the proper maintenance of the perceived upright during changes of posture. The same or some closely related mechanism could conceivably enter into the “assumptive shifts” of the viewer’s position on inspecting figures with ambiguous perspective, or the required reversals of standpoint in dealing with mirror images of the body, where the patient has to go repeatedly from the front to the back view.

General Interpretations

The various symptoms disclosed by our experiments may have to be referred to more than one underlying

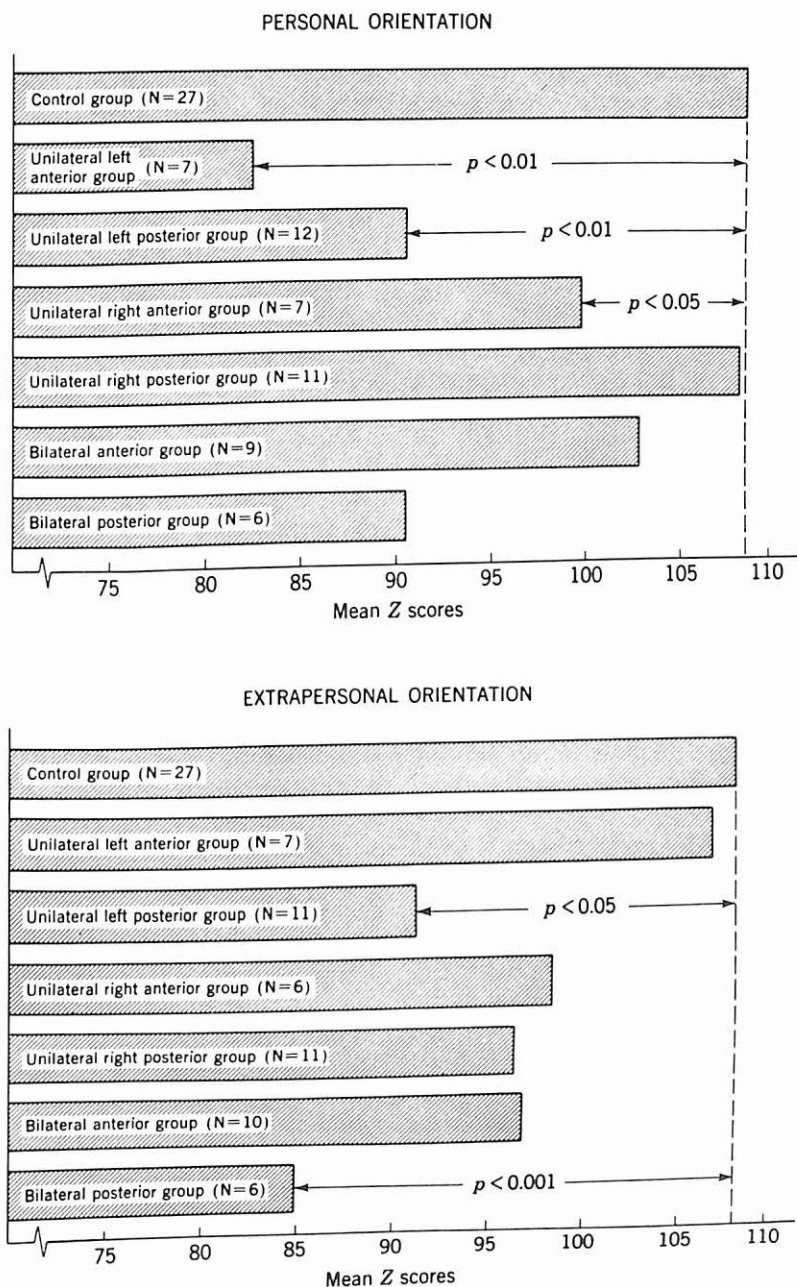


Figure 20.15

Comparison of results (mean Z scores) of tests of personal orientation (see preceding figures) and of extrapersonal orientation; the later test was the visual subtest (top row of maps in Figure 20.12) on the spatial orientation task.

change in function, but we prefer to pursue, for reasons of parsimony at least, and for the time being, a hypothesis of unitary change. As more information is obtained, particularly from cases of sufficiently limited fractional lesions of the frontal lobes, a pattern of symptom dissociation may appear which would conceivably parallel the situation found in the infra-human primate. Nevertheless, the strong "family resemblance" of the symptoms seen in man suggests that we attempt to delineate a common trait in all the components of the syndrome.

The common trait was suggested by perceptual tasks involving posture and movement—either actual movement as in the visual search tests, or virtual movement, as in the other three tasks in which orientation had to be maintained under conditions of changes in posture or perspective. It would be wrong, however, to identify the hypothetical common element in all our frontal symptoms either with control of movement per se, or with regulation of posture, although it is true that the striatum and the overlying frontal cortex can be considered parts of a "supravestibular system" (Muskens, 1934). Posture should not be opposed to movement (a point taken by Denny-Brown, 1962), since movement is always change in posture. Movement as well as posture involve anticipation of change, and an organism has to be able, somehow, to distinguish those changes in sensory input that result from his own, self-produced bodily motions, and those that result from actual movement in the environment. More than half a century ago, von Monakow (1910) surmised that the frontal lobes might play a role in stabilizing perception during changes of posture. We could speculate even

further and propose an "outflow" theory whereby the organism operates upon its own inputs in such a way as to provide continuous distinction between the relative motions of the environment which result from self-produced movement of the body, and the true motions of objects in the surroundings.

We have only to look at the forced movements of the hyperkinetic bifrontal monkey or the curious forced grasping and groping of human frontal lobe disease to be impressed with the need for an analysis of frontal symptoms which stresses the motor element. But it is a particular aspect of motor control that needs to be stressed: it is that aspect which may permit us to give a physiologic meaning to the forbidden concept of "voluntary" movement, in contrast to forced, or reflex, motion.

In this view, every "voluntary" movement has two neural correlates: a stream of impulses to the effectors and a simultaneous "corollary discharge" to central receptor structures, presetting the latter for those predictable changes of input that will be the consequences of the particular motor output. The role of such corollary discharge can be illustrated (as we have done) by considering oculomotor function and frontal eye fields, but this particular neural regulation of eye movement is here proposed as merely one instance of a more general form of sensorimotor coordination which may depend on the integrity of frontal structures and portions of the basal ganglia.

Anatomical pathways for such a postulated central discharge from motor to sensory elements are undoubtedly present. They were known nearly 60 years ago when Probst (1903) described the termination of pyramidal fibers on

dorsal column nuclei. These direct connections from motor to sensory structures made so little sense then, within the framework of stimulus-response physiology, that they were forgotten and had to be rediscovered in more recent studies employing macroelectrodes (Magni et al., 1959), or microelectrodes (Jabbur & Towe, 1961; Towe & Jabbur, 1961), or more conventional histologic techniques (Brodal et al., 1956; Kuypers, 1960; Walberg, 1957). In addition, there is the possible role of a functional "loop" from frontal cortex through the caudate, and various thalamic nuclei, to visual and sensory cortex (Buchwald et al., 1962). As so often happens in such a situation, we do not have too few, but too many anatomical possibilities from which to choose.

Both the pathways and their postulated function are admittedly conjectural, but we are convinced that the assumption of a corollary discharge from anterior to posterior cerebral regions has heuristic value. Without it, the difficulties that cling to the frontal lobe syndromes may be unsurmountable. With it, we may be able to reintroduce the concept of self-produced or voluntary movement, as distinct from automatic reactions: a voluntary motion is any motion that involves a corollary discharge from motor to sensory structures.

Whether the view here presented is right or wrong, the classical analysis of frontal function we believe is doomed to failure. We have to abandon traditional stimulus-response notions. It is not in the reaction to incoming stimuli, but in the prediction of them, the presetting of a mechanism, that the significance of the frontal structures lies. It may in fact be possible to relate this new way of looking at frontal structures to the

earlier studies by Feuchtwanger (1923) which we discussed at such length. His strange-sounding interpretations might be restated in our terms without doing too much violence to his intention.

When he spoke of "will" and "valuation" as the sphere into which the crucial frontal symptoms fall, he expressed strong preference for Ach's introspective accounts of will as an experience. Ach (1905) had said that will involves, subjectively, our "anticipation of the future course of events coupled with an awareness of the role of the self in bringing those events about." Feuchtwanger could have said (but did not) that severe frontal pathology often involves an uncoupling of these two aspects. The patient is not altogether devoid of capacity to anticipate a course of events, but cannot picture himself in relation to those events as a potential agent.

Abnormal fixity, or abnormal oscillation of action, other-directedness, and impulsiveness, all could follow from such a change. We believe that the seemingly lower-level sensorimotor symptoms are made of the same stuff; in fact, they may well be the conditions that give rise, in more severe forms of frontal lobe dysfunction, to the pervasive changes in behavior. These serious changes may derive from the specific symptoms we have found: the inability to give up initial fixations on protracted search, the difficulty in maintaining certain perceptual constancies with varying body posture, the abnormal fluctuations in viewpoint and perspective, and the curious difficulty with shifts of bodily orientation. The interpretation for which we are searching—gropingly, graspingly, as it were—may still elude us, but we believe that there may be continuity between (1) the oculomotor

manifestations (after stimulation or destruction of frontal eye fields), (2) the release of proprioceptive responses from their normal constraint by contact sensations, and (3) the manifold changes in attitude toward more complex tasks, whether set in the laboratory or in everyday life.

Is it too much to hope that the delayed-response deficit, as seen in the bifrontal macaque, might yield to a similar interpretation? Perhaps it is too much to expect that, since it would be compounding a riddle with an enigma. Yet it might be worth trying.

Discussion

Dr. Stamm: I am glad Dr. Teuber brought up again the idea of relating frontal lobes to motor activity. And I was wondering if Dr. Nauta or Dr. Akert have some information about the anatomical connections or circuits which may exist between frontal granular cortex and motor systems.

Dr. Nauta: All the structures known or suspected to receive projections from the prefrontal cortex have a right to be called "motor," if at least that term is not (as it should be) reserved for the motor neurons proper. We must realize that most of the neural structures in the central nervous system are potential links in ascending as well as descending conduction pathways, and therefore

can be considered either "motor" or "sensory," depending on the point of view of the moment. Certainly this is true for the cerebral cortex, even for that part of it referred to as motor cortex. Of course, in its general content your question is well worth asking, and if I may rephrase it, it reads: Where in the brain is the neural outflow of the prefrontal cortex integrated with the activity of other mechanisms involved in the analysis of what went on in the environment? I am afraid that it is yet too early to attempt an answer to this question. The prefrontal cortex appears to have no direct or even oligosynaptic channels to somatic motor neurons, and it can therefore be suspected that its influence upon behavior is exerted rather "high up" in the brain. I know of no reason why the motor cortex should not be considered a possible site of such integration, but so should other structures, and again, let us not be glib in labeling them "motor."

Dr. Pribram: I think that misses the point of Dr. Teuber's beautiful presentation. He said that the feedback has to be to the sensory, not to the motor. And we know that all the "limbic" structures in the

core of the brain deal to a great extent with the activities of the receptive areas. So I do not expect, and would not want from the model that he presented, any more direct input to the motor regions, but I would rather see some mechanism in the core system to preset the sensory ones.

Is that all right with you?

Dr. Teuber: Fine. We share the same delusions, apparently.

Dr. Rosvold: Did I understand you to say, Dr. Teuber, that we don't have a unitary concept of frontal lobe function and we are not likely to develop one?

Dr. Teuber: What I probably said was that I started out by trying to find a unitary concept, but as I moved along, it became rather clear that no single-factor hypothesis could carry one far enough to cover all the manifestations of frontal lesions. And yet the thing that is so tempting to me after this symposium is to think that there may be a family resemblance among symptoms, even among those which seem in part dissociable. This may have anatomical corollaries in that the two parts that we now think of as having somewhat divergent functions, dorsal lateral and orbital, are not easily dissociated at lower lev-

els to which they project or from which they receive input; they are always very close together. Any lesion, therefore, would tend to straddle the dividing line between them, and the symptoms are not that dissimilar. Even though Dr. Mishkin got some dissociation, it looked more like single dissociation rather than double dissociation, for the orbital as compared with the dorsal lateral syndrome.

Dr. Konorski: Dr. Teuber, you know I am in strong agreement with you as far as the value of speculation is concerned, therefore, I would like to speculate and propose some explanations for the four phenomena which you described in your talk.

The first, the searching for the figure, could be explained by the notion of "stimulus boundness," that the patient with frontal lobe lesion is so very much bound to this central figure that it is really difficult for him to switch to other figures, and hence this increased latency. Perhaps the following experiment would prove illuminating as to the basis of the frontal's loss; remove the central model figure from the slide when the patient is required to search for the proper forms so that

there is no longer a stimulus to which he is bound, and his performance might improve in consequence.

As to the problem of your reversible figure test, I think that this switching of the figure has something to do with eye movements and with functions of area 8, in other words inability or decreased ability for the "visual grasping" of an object, instead, they switch from one object to the other in a nystagmic manner.

As to the fourth test, namely, bodily disorientation with respect to the mirror image, I think that this is a very complicated task, difficult to deal with even for normals. Perhaps the difficulty in performing this task has to do with inflexibility, or difficulty in shifting. The same process was named differently by different participants of this symposium. Certainly, this is the sort of task which requires flexibility.

And now the third test! I think that you worked so much on it that I do not dare to offer a comment.

(Laughter)

You have certainly exhausted the stock of available explanations for this one.

Dr. Teuber:

I think I could agree with you on every point, except perhaps the question of the role of overt eye movements, but with the important qualification that all these terms are purely descriptive of the disability we see in frontal cases, such as inability to shift, inflexibility, rigidity, hyperfixation, merely referring to *capacity* or lack of it. I am groping, and that is all I am doing, for a description in terms of *process*. I think the description in terms of capacity is almost a faculty description, you see. It is a good shorthand to designate the symptom, but no more. I would like to go beyond this, to identify the underlying functional change. I am very much aware of the fact that I haven't got there, but I would like to try.

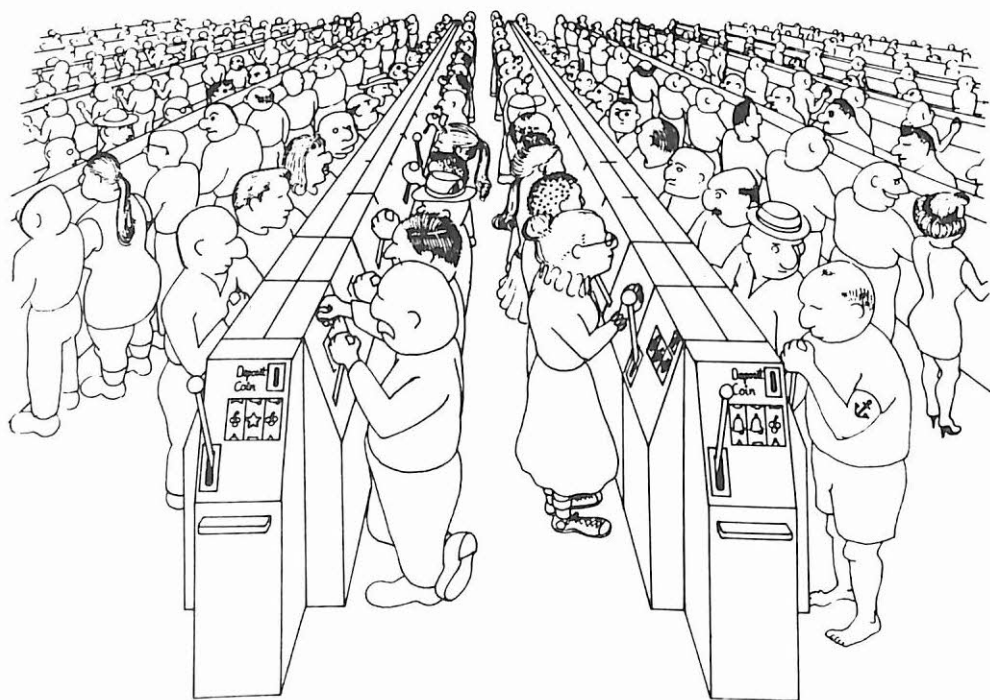
To go back quickly, I think it is quite apparent we must now reexamine the field-of-search test, under two conditions: (1) with fixation maintained, and (2) while permitting the patient's gaze to roam freely over the array; this should be done in order to see whether he is perhaps paradoxically improved by such a permission not to fixate. Also, I think we should record the

complete search process, because there may be very peculiar kinds of search going on in these people.

Dr. Konorski raised the question of abnormal eye movement as a basis for the abnormal response to reversible figures. However, abnormalities of overt eye movements have been ruled out fairly effectively, at least insofar as they could be the direct cause of these shifts. Recordings of eye movements during prolonged inspection of such reversible cubes do not suggest that the eye movements in any way precede a given shift. Sometimes there is a flick after a

shift, but not before. As I said, the basic trouble may lie in diminished capacity for presetting the sensory systems, in anticipation of reafferent changes consequent upon motion. In that sense, the trouble involves the oculomotor system, but it does so on a central level.

It is still possible that the personal orientation test may be reflecting analogous central difficulties with reversal of set. It may be of some comfort to Dr. Milner to know that we do have a strong positive correlation between poor performance on this task and failure on the Wisconsin Card Sorting Test.



PART FIVE

control of human behavior

Perhaps the most central question in psychology is "What are the determinants of human behavior?" One of Freud's outstanding contributions was his emphasis on the importance of factors of which individuals may not be aware or conscious. The first selection in this group gives some of Freud's comments on slips of the tongue. The second article is concerned with a vital problem for society: providing children with a fruitful and productive experience in school. Hamblin and his co-authors have studied the effects of simple "contingencies of reinforcement" on constructive and disruptive classroom behavior. The results are most convincing. Next, Goldiamond discusses some simple procedures which have been used to enhance self-control—an area of immense interest which psychologists have been wont to overlook. And the final selection in this group, by B. F. Skinner, continues that important discussion. It has been argued that people are singularly incompetent at describing the causative agents responsible for their own behavior. As you read these papers, try to decide whether you agree with this viewpoint.

INTRODUCTION

Many of Freud's contributions to our understanding of human behavior give us insights into the complex nature of human motivation. In the selection that follows, three discrete levels of motivation are discussed.

- 1) Conscious motivation overtly expressed: the situation in which a person understands his own purposes and publicly expresses them.
- 2) Conscious motivation without public expression: the situation in which a person is perfectly aware of why he is doing something or what he is thinking but keeps this awareness to himself.
- 3) Unconscious motivation: the situation in which a person acts without awareness of the wellsprings of his action.

We can see all three kinds of motivation in various parts of the section which follows. Here, Freud concentrated his attention on slips of the tongue, those innocuous little verbal errors which reveal an underlying motivation pattern a) of which we were totally unaware or b) of which we were aware but which we wished to keep secret or c) of which we were aware but which seemed totally irrelevant to the present discussion. All three of these kinds of intrusions appear in various of the slips of the tongue discussed below. By looking at the section as a whole, we can see some striking evidence of the complex nature of the determinants of human speech.

This excerpt is taken from Freud's fascinating and highly readable book *Psychopathology of Everyday Life*. If you enjoy this fragment, you will surely like the whole book.

SLIPS OF THE TONGUE

S. Freud

* A slip of the tongue [occurred] in the case of another woman patient, whose memory failed her in the middle of reproducing a long-lost recollection of childhood. Her memory would not tell her what part of her body had been grasped by a prying and lascivious hand. Immediately afterwards she called on a friend with whom she discussed summer residences. When she was asked where her cottage at M. was situated she answered: 'on the *Berg-lende* [hill-thigh] instead of *Berglehne* [hill-side].

* At a certain point in the analysis of another woman patient I had to tell her that I suspected her of having been ashamed of her family during the period we were just then concerned with, and of having reproached her father with something we did not yet know about. She remembered nothing of the kind and moreover declared it was unlikely. However, she continued the conversation with some remarks about her family: 'One thing must be granted them: they are certainly unusual people, they all possess *Geiz* [greed]—I meant to say "*Geist*

[cleverness]".'

 And this was in fact the reproach which she had repressed from her memory. It is a frequent occurrence for the idea one wants to withhold to be precisely the one which forces its way through in the form of a slip of the tongue.

* A young man said to his sister: 'I've completely fallen out with the D.'s now. We're not on speaking terms any longer.' 'Yes indeed!' she answered, 'they're a fine *Lippschaft*.' She meant to say '*Sippschaft* [lot, crew]', but in the slip she compressed two ideas: viz. that her brother had himself once begun a flirtation with the daughter of this family, and that this daughter was said to have recently become involved in a serious and irregular *Liebschaft* [love-affair].

* A young man addressed a lady in the street in the following words: 'If you will permit me, madam, I should like to "*begleit-digen*" you.' It was obvious what his thoughts were: he would like to '*begleiten* [accompany]' her, but was afraid his offer would '*beleidigen* [insult]' her. That these two conflicting emotional impulses found expression in one word—in the slip of the tongue, in fact—indicates that the young man's real intentions were at any rate not of the purest, and were bound to seem, even to himself, insulting to the lady. But while he attempted to conceal this from her, his unconscious played a trick on him by

Reprinted from *The Psychopathology of Everyday Life* by Sigmund Freud. Translated from the German by Alan Tyson. Edited with an introduction and additional notes by James Strachey. By permission of W. W. Norton & Co., Inc., and Ernest Benn Ltd. Editorial matter Copyright © 1960, 1965 by James Strachey. Translation Copyright © 1960 by Alan Tyson.

betraying his real intentions. But on the other hand he in this way, as it were, anticipated the lady's conventional retort: 'Really! What do you take me for? How dare you *insult* me!' (Reported by O. Rank.)

I will next quote a number of examples from an article by Stekel, entitled 'Unconscious Admissions', in the *Berliner Tageblatt* of January 4, 1904.

* 'An unpleasant part of my unconscious thoughts is disclosed by the following example. I may start by stating that in my capacity as a doctor I never consider my remuneration but only have the patient's interest in mind: that goes without saying. I was with a woman patient to whom I was giving medical attention in a period of convalescence after a serious illness. We had been through hard days and nights together. I was happy to find her improved; I painted a picture for her benefit of the delights of a stay in Abbazia, and concluded by saying: "If, as I hope, you will *not* leave your bed soon. . . ." This obviously owed its origin to an egoistic motive in the unconscious, namely that I should be able to continue treating this well-to-do patient some time longer—a wish that is entirely foreign to my waking consciousness and which I would indignantly repudiate.'

* Here is another example from Stekel. 'My wife was engaging a French governess for the afternoons, and after agreement had been reached on the terms, wanted to retain her testimonials. The Frenchwoman asked to be allowed to keep them, giving as her reason: *Je cherche encore pour les après-midis, pardon, pour les avant-midis* [I am still looking for work in the afternoons—I mean, in the forenoons]. She obviously had the intention of looking round elsewhere and

perhaps finding better terms—an intention which she in fact carried out.'

* From Stekel: 'I had to give a stiff lecture to a wife; and her husband, at whose request I did it, stood outside the door listening. At the end of my sermon, which had made a visible impression, I said: "Good-bye, sir." To any well-informed person I was thus betraying the fact that my words were addressed to the husband and that I had spoken them for his benefit.'

* Stekel reports of himself that at one time he had two patients from Trieste in treatment whom he always used to address the wrong way round. 'Good morning, Herr Peloni', he would say to Askoli, and 'Good morning, Herr Askoli' to Peloni. He was at first inclined not to attribute any deeper motive to this confusion but to explain it as being due to the numerous points of resemblance between the two gentlemen. However it was easy for him to convince himself that the interchanging of the names corresponded in this case to a kind of boastfulness: he was able in this way to let each of his Italian patients know that he was not the only visitor from Trieste who had come to Vienna in search of his medical advice.

* Stekel reports that during a stormy General Meeting he said: 'We shall now *streiten* [quarrel]' (instead of '*schreiten* [proceed]') 'to point four on the agenda.'

* A professor declared in his inaugural lecture: 'I am not *geneigt* [inclined]' (instead of '*geeignet* [qualified]') 'to describe the services of my most esteemed predecessor.'

* Occasionally a slip of the tongue takes the place of a detailed characterization. A young woman who wore the breeches in her home told me that her sick husband had been to the doctor

to ask what diet he ought to follow for his health. The doctor, however, had said that a special diet was not important. She added: 'He can eat and drink what *I* want.'

The following two examples by Reik (1915) have their origin in situations where slips of the tongue occur especially easily—situations in which more must be kept back than can be said.

* A gentleman was offering his condolences to a young lady whose husband had recently died, and he intended to add: 'You will find consolation in *devoting* [*widmen*] yourself entirely to your children.' Instead he said '*widwen*'. The suppressed thought referred to consolation of another kind: a young and pretty widow [*Witwe*] will soon enjoy fresh sexual pleasures.

* At an evening party the same gentleman was having a conversation with the same lady about the extensive preparations being made in Berlin for Easter, and asked: 'Have you seen today's display [*Auslage*] at Wertheim's? The place is completely *decollated*.' He had not dared to express his admiration for the beautiful lady's *décolletage*, while the word '*Auslage*' ['display'] was used unconsciously in two senses.

The same condition applies to another case, observed by Dr. Hanns Sachs, of which he has tried to give an exhaustive account:

* 'A lady was telling me about a common acquaintance. The last time she saw him, he was, she said, as elegantly dressed as ever: in particular he was wearing strikingly beautiful brown *Halbschuhe* [low shoes]. When I asked where she had met him she replied: "He rang at the door of my house and I saw him through the blinds, which were down. But I didn't open the door or give any other sign of life, as I

didn't want him to know I was already back in town." While I was listening to her I had an idea that she was concealing something from me, most probably the fact that her reason for not opening the door was that she was not alone and not properly dressed to receive visitors; and I asked her somewhat ironically: "So you were able to admire his *Hausschuhe* [house shoes]—*Halbschuhe* [low shoes], I mean—through the blinds when they were drawn?" In *Hausschuhe* I was giving expression to the thought of her *Hauskleid* [lit. house dress, i.e. *négligée*] which I had refrained from uttering. There was on the other hand a temptation to get rid of the word "*Halb* [half]" for the reason that it was precisely this word which contained the core of the forbidden answer: "You are only telling me *half* the truth and are hiding the fact that you were *half* dressed." The slip of the tongue was encouraged by the additional circumstance that we had been talking directly before about this particular gentleman's married life, about his *häuslich* [domestic] happiness; this no doubt helped to determine the displacement [of "*Haus*"] on to him. Finally, I must confess that envy on my part may perhaps have contributed to my placing this elegant gentleman in the street in house shoes; only recently I myself bought a pair of brown low shoes, which are certainly not "strikingly beautiful" any longer.'

Times of war like the present produce numerous slips of the tongue which there is not much difficulty in understanding.

* 'What regiment is your son with?' a lady was asked. She replied: 'With the 42nd Murderers' [*Mörder*—instead of '*Mörser*', 'Mortars'].

* Here is a quite especially instructive slip of the tongue which I should

not like to omit, although according to my authority it is some twenty years old. 'A lady once advanced the following opinion at a social gathering—and the words show that they were uttered with fervour and under the pressure of a host of secret impulses: "Yes, a woman must be pretty if she is to please men. A man is much better off; as long as he has his *five* straight [*fünf gerade*] limbs he needs nothing more!" This example allows us a good view of the intimate mechanism of a slip of the tongue that results from *condensation* or from a *contamination*. It is plausible to suppose that we have here a fusion of two turns of phrase with similar meanings:

as long as he has his *four straight limbs*
as long as he has his *five wits* about
him.

Or the element *straight* [*gerade*] may have been common to two intended expressions which ran:

as long as he has his *four straight limbs*
to treat all *five(s)* as *even numbers*¹

'There is nothing in fact to prevent us from assuming that *both* turns of phrase, the one about his five wits and the one about 'the even number five', played their separate parts in causing first a number, and then the mysterious five instead of the simple four, to be introduced into the sentence deal-

¹ [*Alle fünf gerade sein lassen.* The German '*gerade*' means both 'straight' and 'even'. The meaning of the phrase, literally translated in the text, is: 'To close one's eyes to irregularities.']

ing with the straight limbs. But this fusion would certainly not have come about if, in the form that appeared in the slip of the tongue, it had not had a good meaning of its own—one expressing a cynical truth which could not of course be admitted to undisguised, coming as it did from a woman. —Finally we should not omit to draw attention to the fact that the lady's remark, as worded, could pass just as well for a capital joke as for an amusing slip of the tongue. It is simply a question of whether she spoke the words with a conscious or an unconscious intention. In our case the way the speaker behaved certainly ruled out any notion of conscious intention and excluded the idea of its being a joke.'

* At this point I insert another slip of the tongue which it needs little skill to interpret. 'In an anatomy lesson the professor was endeavouring to explain the nasal cavities, which are notoriously a very difficult department of enterology. When he asked whether his audience had understood his presentation of the subject, he received a general reply in the affirmative. Thereupon the professor, who was known for his high opinion of himself, commented: 'I can hardly believe that, since, even in Vienna with its millions of inhabitants, those who understand the nasal cavities can be counted *on one finger*, I mean on the fingers of one hand.'

* On another occasion the same professor said: 'In the case of the female genitals, in spite of many *Versuchungen* [temptations]—I beg your pardon, *Versuche* [experiments] . . .'

INTRODUCTION

This article, written by a group of sociologists and educators, describes the use of what psychologists refer to as "behavior modification" techniques. The basic assumption on which these techniques are founded is so simple as to appear almost laughable: reward the behavior which you wish to occur and make sure that you don't reward a behavior which you wish to eliminate. Simple as this injunction appears, it is far less simple to apply it effectively. To do so you must have completely clear criteria to determine when behavior of various kinds is occurring and you must know precisely what constitutes a reward for the people you are working with.

Basically, the authors are concerned with how you get "problem" children to learn. Starting with the simple approach mentioned above, and outlined in far greater detail in the article, the authors have demonstrated remarkable progress. You may wish to note the following points which are amplified in the paper:

1. Children respond to the contingencies which are in effect—i.e., the conditions under which they obtain rewards. Children do not respond to the wishes of adults. With the best intentions in the world, many adults inadvertently reinforce or reward the very behavior which they are trying to suppress. Several examples of this sort are described below.
2. In the experiment on hyperaggressive children, note that the authors did not try simply to eliminate aggression. They recognized that in order to do so they must also provide a new non-aggressive activity for the child. Thus, at the same time they were trying to eliminate aggression, they tried to divert the child's activities in a constructive direction by reinforcing him for his cooperation. Either half of this program might have failed in isolation but the two halves together—rewarding cooperation while ignoring aggression—did have a remarkable effect.
3. Note that at no point in the experiments described was any deliberate punishment used except for withholding a reward.
4. Note that for the very aggressive children even being *fined* a token appeared to be a reinforcer. That is, although they were losing a desirable token, the fact that the teacher approached them and interacted with them apparently was attractive enough to overcome the aversive com-

ponent of losing the token. Here is another example of a case in which the intention of the adult to punish behavior (by means of a fine) was thwarted because the reinforcer used turned out to be a reward rather than an aversive consequence.

5. In the studies on cooperation, we can expect that the results might have been even more impressive if the experimenters had shifted to a schedule of partial reinforcement of cooperative responding, i.e., if they had only reinforced occasional cooperative responses rather than reinforcing every one.

6. Note in the experiment, results of which are presented in Figure 6, that after the token reinforcements were removed late in the series, the children continued to talk almost as much as with reinforcement. This information is of vital importance since it means that after the therapeutic sessions were completed the normal rewards for talking, available to everyone, were sufficient to maintain talking in those children who, before, rarely spoke to strangers. Such a situation means that the treatment is in a real sense a therapy since it can be applied until the cure is affected and then removed, leaving the child in a "healthier" state.

Do not be misled by the remarkably simple notion underlying this article. Simple as the basic idea appears, and obvious though it may have been to your grandmother and ours, it is nonetheless extraordinarily potent and surprisingly difficult to implement effectively.

CHANGING THE GAME FROM "GET THE TEACHER" TO "LEARN"

*Robert L. Hamblin / David Buckholdt / Donald Bushell /
Desmond Ellis / Daniel Ferritor*

Almost any educator of experience will assure you that it is next to impossible—and often actually impossible—to teach normal classroom subjects to children who have extreme be-

havior problems, or who are "too young." Yet at four experimental classrooms of the Central Midwestern Regional Educational Laboratories (CEMREL), we have been bringing about striking changes in the behavior and learning progress of just such children.

Reprinted with permission from Trans-action, 6, 1969, pp. 20-31. Copyright © 1969 by Trans-action magazine, St. Louis, Missouri.

In the 18 months of using new exchange systems and working with different types of problem children, we have seen these results:

■ Extraordinarily aggressive boys, who had not previously responded to therapy, have been tamed.

■ Two-year-olds have learned to read about as fast and as well as their 5-year-old classmates.

■ Four ghetto children, too shy, too withdrawn to talk, have become better than average talkers.

■ Several autistic children, who were either mute or could only parrot sounds, have developed functional speech, have lost their bizarre and disruptive behavior patterns, and their relationships with parents and other children have improved. All of these children are on the road to normality.

Our system is deceptively simple. Superficially, in fact, it may not even seem new—though, in detail, it has never been tried in precisely this form in the classroom before. In essence, we simply reinforce “good” behavior and nonpunitively discourage “bad” behavior. We structure a social exchange so that as the child progresses, we reinforce this behavior—give him something that he values, something that shows our approval. Therefore, he becomes strongly motivated to continue his progress. To terminate bizarre, disruptive or explosive patterns, we stop whatever has been reinforcing that undesirable behavior—actions or attention that teachers or parents have unwittingly been giving him in exchange, often in the belief that they were punishing and thus discouraging him. Study after study has shown that whenever a child persists in behaving badly, some adult has, perhaps inadvertently, been rewarding him for it. “Socialization” is the term that so-

ciologists use to describe the process of transforming babies—who can do little but cry, eat, and sleep—into adults who can communicate and function rather effectively in their society. Socialization varies from culture to culture, and, while it is going on all around us, we are seldom aware of it. But when normal socialization breaks down, “problems” occur—autism, non-verbal or hyperaggressive behavior, retardation, delinquency, crime, and so on.

The authors, after years of often interesting but by and large frustrating research, realized that the more common theories of child development (Freudian, neo-Freudian, the developmental theories of Gesell and Piaget, and a number of others) simply do not satisfactorily explain the socialization process in children. Consequently in desperation we began to move toward the learning theories and then toward the related exchange theories of social structure. Since then, working with problem children, our view has gradually been amplified and refined. Each experimental classroom has given us a different looking glass. In each we can see the child in different conditions, and can alter the conditions which hinder his socialization into a civilized, productive adult capable of happiness.

By the time they become students, most children love to play with one another, to do art work, to cut and paste, to play with Playdoh, to climb and swing on the playground, and so on. Most pre-schools also serve juice and cookie snacks, and some have television sets or movies. There is, consequently, no dearth of prizes for us to reward the children for good behavior. The problem is not in finding reinforcers, but in managing them.

The Basic System: Token Exchange

One of the simpler and most effective ways, we found, was to develop a token-exchange system. The tokens we use are plastic discs that children can earn. A child who completes his arithmetic or reading may earn a dozen tokens, given one by one as he proceeds through the lessons. And at the end of the lesson period comes the reward.

Often it is a movie. The price varies. For four tokens, a student can watch while sitting on the floor; for eight, he gets a chair; for 12, he can watch while sitting on the table. Perhaps the view is better from the table—anyway, the children almost always buy it if they have enough tokens. But if they dawdled so much that they earned fewer than four, they are "timed out" into the hall while the others see the movie. Throughout the morning, therefore, the children earn, then spend, then earn, then spend.

This token-exchange system is very powerful. It can create beneficial changes in a child's behavior, his emotional reactions, and ultimately even his approach to life. But it is not easy to set up, nor simple to maintain.

At the beginning the tokens are meaningless to the children; so to make them meaningful, we pair them with M&M candies, or something similar. As the child engages in the desired behavior (or a reasonable facsimile), the teacher gives him a "Thank you," an M&M, and a token. At first the children are motivated by the M&Ms and have to be urged to hold on to the tokens; but then they find that the tokens can be used to buy admission to the movie, Playdoh, or other good things. The teacher tells

them the price and asks them to count out the tokens. Increasingly, the teacher "forgets" the M&Ms. In two or three days the children get no candy, just the approval and the tokens. By then, they have learned.

There are problems in maintaining a token exchange. Children become disinterested in certain reinforcers if they are used too frequently, and therefore in the tokens that buy them. For instance, young children will work very hard to save up tokens to play with Playdoh once a week; if they are offered Playdoh every day, the charm quickly fades. Some activities—snacks, movies, walks outdoors—are powerful enough to be used every day.

As noted, the children we worked with had different behavior problems, reflecting various kinds of breakdowns in the socialization process. Each experiment we conducted concentrated on a particular type of maladjustment or a particular group of maladjusted children to see how a properly structured exchange system might help them. Let us look at each experiment, to see how each problem was affected.

Aggression

Unfortunately, our world reinforces and rewards aggressive behavior. Some cultures and some families are open and brazen about it—they systematically and consciously teach their young that it is desirable, and even virtuous, to attack certain other individuals or groups. The child who can beat up the other kids on the playground is sometimes respected by his peers, and perhaps by his parents; the soldier achieves glory in combat. The status, the booty, or the bargaining advantages that come to the aggressor

can become reinforcement to continue and escalate his aggressions.

In more civilized cultures the young are taught not to use aggression, and we try to substitute less harmful patterns. But even so, aggression is sometimes reinforced unintentionally—and the consequences, predictably, are the same as if the teaching was deliberate.

In the long run civilized cultures are not kind to hyperaggressive children. A recent survey in England, for instance, found that the great majority of teachers felt that aggressive behavior by students disturbed more classrooms than anything else and caused the most anxiety among teachers. At least partly as a result, the dropout rates for the hyperaggressives was $2\frac{1}{2}$ times as great as for "normals," and disproportionate numbers of hyperaggressives turned up in mental clinics.

The traditional treatment for aggressive juveniles is punishment—often harsh punishment. This is not only of dubious moral value, but generally it does not work.

We took seriously—perhaps for the first time—the theory that aggression is a type of exchange behavior. Boys become aggressive because they get something for it; they continue to be aggressive because the rewards are continuing. To change an aggressive pattern in our experimental class at Washington University, therefore, we had to restructure appropriately the exchange system in which the boys were involved.

As subjects we (Ellis and Hamblin) found five extraordinarily aggressive 4-year-old boys, all referred to us by local psychiatrists and social workers who had been able to do very little with them. Next, we hired a trained teacher. We told her about the boys and the general nature of the experi-

ment—then gave her her head. That is, she was allowed to use her previous training during the first period—and this would provide a baseline comparison with what followed after. We hoped she would act like the "typical teacher." We suspect that she did.

Let's Play "Get the Teacher"

The teacher was, variously, a strict disciplinarian, wise counselor, clever arbitrator, and sweet peacemaker. Each role failed miserably. After the eighth day, the average of the children was 150 sequences of aggression per day! Here is what a mere four minutes of those sequences were like:

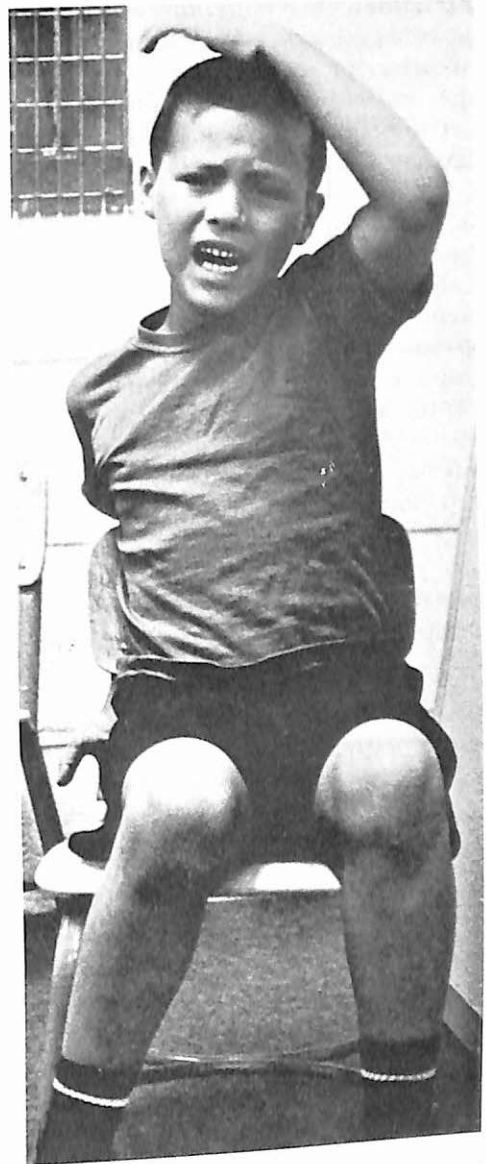
Mike, John, and Dan are seated together playing with pieces of Playdoh. Barry, some distance from the others, is seated and also is playing with Playdoh. The children, except Barry, are talking about what they are making. Time is 9:10 A.M. Miss Sally, the teacher, turns toward the children and says, "It's time for a lesson. Put your Playdoh away." Mike says, "Not me." John says, "Not me." Dan says, "Not me." Miss Sally moves toward Mike. Mike throws some Playdoh in Miss Sally's face. Miss Sally jerks back, then moves forward rapidly and snatches Playdoh from Mike. Puts Playdoh in her pocket. Mike screams for Playdoh, says he wants to play with it. Mike moves toward Miss Sally and attempts to snatch the Playdoh from Miss Sally's pocket. Miss Sally pushes him away. Mike kicks Miss Sally on the leg. Kicks her again, and demands the return of his Playdoh. Kicks Miss Sally again. Picks up a small steel chair and throws it at Miss Sally. Miss Sally jumps out of the way. Mike picks up another chair and throws it more violently. Miss Sally cannot move in time. Chair strikes her foot. Miss Sally pushes Mike down on the floor. Mike starts up. Pulls

over one chair. Now another, another. Stops a moment. Miss Sally is picking up chairs, Mike looks at Miss Sally. Miss Sally moves toward Mike. Mike runs away.

John wants his Playdoh. Miss Sally says "No." He joins Mike in pulling over chairs and attempts to grab Playdoh from Miss Sally's pocket. Miss Sally pushes him away roughly. John is screaming that he wants to play with his Playdoh. Moves toward phonograph. Pulls it off the table; lets it crash onto the floor. Mike has his coat on. Says he is going home. Miss Sally asks Dan to bolt the door. Dan gets to the door at the same time as Mike. Mike hits Dan in the face. Dan's nose is bleeding. Miss Sally walks over to Dan, turns to the others, and says that she is taking Dan to the washroom and that while she is away, they may play with the Playdoh. Returns Playdoh from pocket to Mike and John. Time: 9:14 A.M.

Wild? Very. These were barbarous little boys who enjoyed battle. Miss Sally did her best but they were just more clever than she, and they *always* won. Whether Miss Sally wanted to or not, they could always drag her into the fray, and just go at it harder and harder until she capitulated. She was finally driven to their level, trading a kick for a kick and a spit in the face for a spit in the face.

What Miss Sally did not realize is that she had inadvertently structured an exchange where she consistently reinforced aggression. First, as noted, whenever she fought with them, she *always lost*. Second, more subtly, she reinforced their aggressive pattern by giving it serious attention—by looking, talking, scolding, cajoling, becoming angry, even striking back. These boys were playing a teasing game called "Get the Teacher." The more she showed that she was bothered by their



behavior, the better they liked it, and the further they went.

These interpretations may seem far-fetched, but they are borne out dramatically by what happened later. On the twelfth day we changed the conditions, beginning with B1 (see Figure 1). First, we set up the usual token exchange to reinforce cooperative be-

havior. This was to develop or strengthen behavior that would replace aggression. Any strong pattern of behavior serves some function for the individual, so the first step in getting rid of a strong, disruptive pattern is substituting another one that is more useful and causes fewer problems. Not only therapy, but simple humanity dictates this.

First, the teacher had to be instructed in how *not to reinforce* aggression. Contrary to all her experience, she was asked to turn her back on the aggressor, and at the same time to reinforce others' cooperation with tokens. Once we were able to coach her and give her immediate feedback over a wireless-communication system, she structured the exchanges almost perfectly. The data in Figures 1 and 2 show the crucial changes: a gradual

increase in cooperation—from about 56 to about 115 sequences per day, and a corresponding decrease in aggression from 150 to about 60 sequences!

These results should have been satisfactory, but we were new at this kind of experimentation, and nervous. We wanted to reduce the frequency of aggression to a "normal" level, to about 15 sequences a day. So we restructured the exchange system and thus launched A2.

In A2, we simply made sure that aggression would always be punished. The teacher was told to *charge* tokens for any aggression.

To our surprise, the frequency of cooperation remained stable, about 115 sequences per day; but aggression *increased* to about 110 sequences per day! Evidently the boys were still playing "Get the Teacher," and the fines

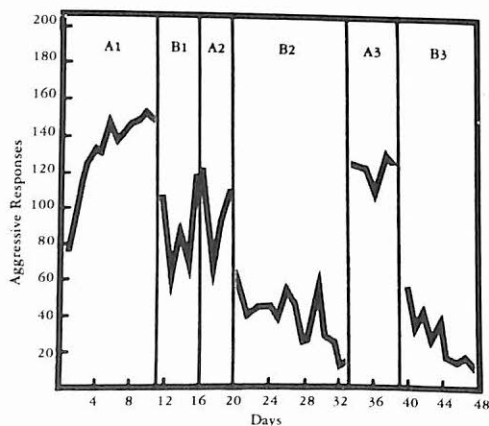


Figure 1

Frequency of aggressive sequences by days for five 4-year-old boys. In A1, A2 and A3 the teacher attempted to punish aggression but inadvertently reinforced it. In B1, B2 and B3 she turned her back or otherwise ignored aggression and thus did not reinforce it.

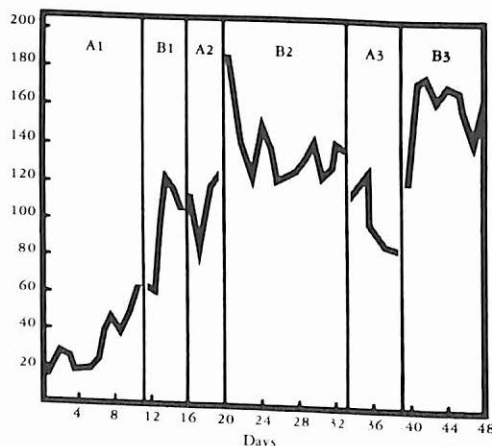


Figure 2

Frequency of cooperative sequences. In A1, A2 and A3 the teacher structured a weak approval exchange for cooperation and a disapproval exchange for noncooperation. In B1, A2, B2 and B3, she structured a token exchange for cooperation.

were enough reinforcement to increase aggression.

So, instead of fining the children, the teacher was again told to ignore aggression by turning her back and giving attention and tokens only for cooperation. The frequency of aggression went down to a near "normal" level, about 16 sequences per day (B2), and cooperation increased to about 140 sequences.

Then, as originally planned, the conditions were again reversed. The boys were given enough tokens at the beginning of the morning to buy their usual supply of movies, toys, and snacks, and these were not used as reinforcers. The teacher was told to do the best she could. She was not instructed to return to her old pattern, but without the tokens and without our coaching she did—and with the same results. Note A3 in Figures 1 and 2. Aggression increased to about 120 sequences per day, and cooperation decreased to about 90. While this was an improvement over A1, before the boys had ever been exposed to the token exchange, it was not good. The mixture of aggression and cooperation was strange, even weird, to watch.

When the token exchange was restructured (B3) and the aggression no longer reinforced, the expected changes recurred—with a bang. Aggression decreased to seven sequences on the last day, and cooperation rose to about 181 sequences. In "normal" nursery schools, our observations have shown that five boys can be expected to have 15 aggression sequences and 60 cooperation sequences per day. Thus, from extremely aggressive and uncooperative, our boys had become less aggressive and far more cooperative than "normal" boys.

Here is an example of their new

behavior patterns, taken from a rest period—precisely the time when the most aggressive acts had occurred in the past:

All of the children are sitting around the table drinking their milk; John, as usual, has finished first. Takes his plastic mug and returns it to the table. Miss Martha, the assistant teacher, gives him a token. John goes to cupboard, takes out his mat, spreads it out by the blackboard, and lies down. Miss Martha gives him a token. Meanwhile, Mike, Barry, and Jack have spread their mats on the carpet. Dan is lying on the carpet itself since he hasn't a mat. Each of them gets a token. Mike asks if he can sleep by the wall. Miss Sally says "Yes." John asks if he can put out the light. Miss Sally says to wait until Barry has his mat spread properly. Dan asks Mike if he can share his mat with him. Mike says "No." Dan then asks Jack. Jack says, "Yes," but before he can move over, Mike says "Yes." Dan joins Mike. Both Jack and Mike get tokens. Mike and Jack get up to put their tokens in their cans. Return to their mats. Miss Sally asks John to put out the light. John does so. Miss Martha gives him a token. All quiet now. Four minutes later—all quiet. Quiet still, three minutes later. Time: 10:23 A.M. Rest period ends.

The hyperaggressive boys actually had, and were, double problems; they were not only extremely disruptive, but they were also washouts as students. Before the token system (A1), they paid attention to their teacher only about 8 percent of the lesson time (see Figure 3). The teacher's system of scolding the youngsters for inattention and taking their attention for granted with faint approval, if any, did not work at all. To the pupils, the "Get the Teacher" game was much more satisfying.

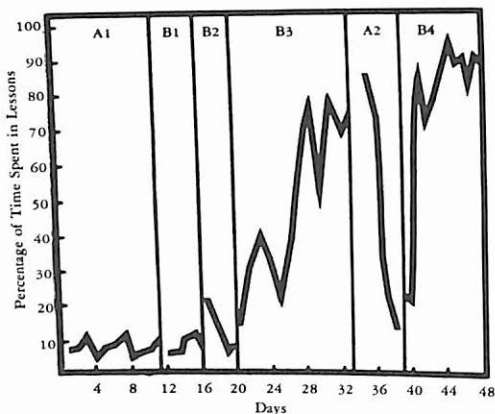


Figure 3

Percentage of scheduled time spent in lessons by days for five hyperaggressive boys. In A1 and A2, teacher structured approval exchange for attendance, disapproval for non-attendance. In B1 and B2, a token exchange for attendance was structured, but not effectively until B2 and B4.

After the token exchange was started, in B1, B2, B3, and B4, it took a long, long time before there was any appreciable effect. The teacher was being trained from scratch, and our methods were, then, not very good. However, after we set up a wireless-communication system that allowed us to coach the teacher from behind a one-way mirror and to give her immediate feedback, the children's attention began to increase. Toward the end of B3, it leveled off at about 75 percent—from 8 percent! After the token exchange was taken out during A2, attention went down to level off at 23 percent; put back in at B4, it shot back up to a plateau of about 93 percent. Like a roller coaster: 8 percent without, to 75 with, to 23 without, to 93 with.

Normal Children

These results occurred with chronic, apparently hopeless hyperaggressive boys. Would token exchange also help "normal," relatively bright upper-middle-class children? Sixteen youngsters of that description—nine boys and seven girls, ranging from 2 years 9 months to 4 years 9 months—were put through an experimental series by Bushell, Hamblin, and Denis Stoddard in an experimental pre-school at Webster College. All had about a month's earlier experience with the token-exchange system. The results are shown in Figure 4.

Study in 15-Minute Periods

At first, the study hour was broken up into 15-minute periods, alternating between the work that received tokens, and the play or reward that the tokens could be used for. Probably because the children were already familiar with token exchange, no great increase in learning took place. On the 22nd day, we decided to try to increase the learning period, perhaps for the whole hour. In A2 (Figure 4), note that the time spent in studying went up rapidly and dramatically—almost doubling—from 27 to level off at 42 minutes.

During B, the token exchange was taken out completely. The teachers still gave encouragement and prepared interesting lessons as before. The rewards—the nature walks, snacks, movies, and so on—were retained. But, as in a usual classroom, they were given to the children free instead of being sold. The children continued at about the same rate as before for a few days. But after a week, attention dropped off slowly, then sharply.

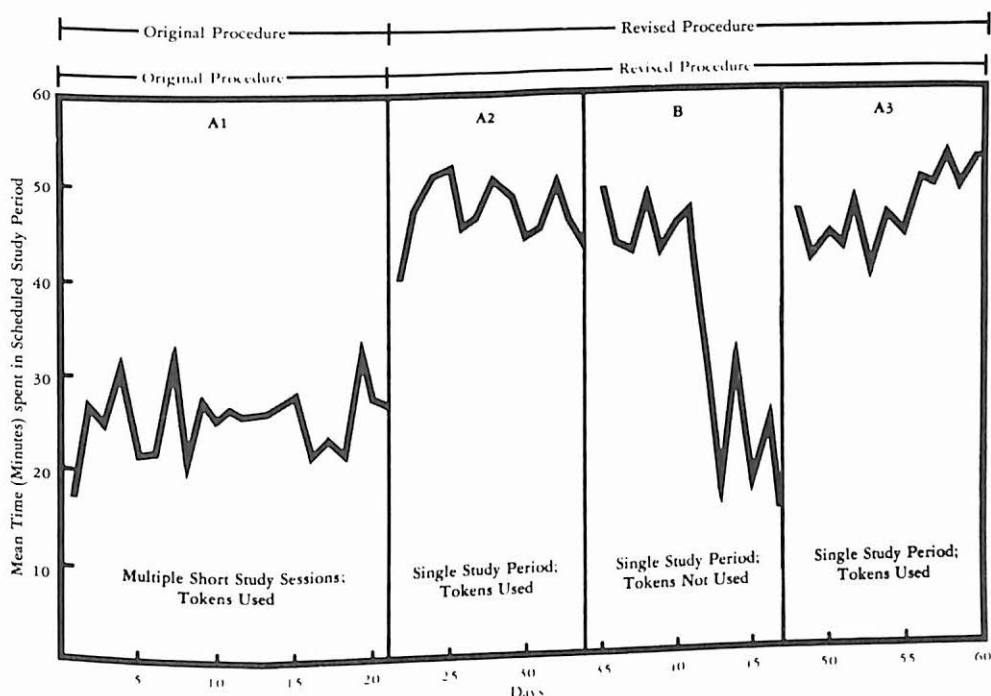


Figure 4

Sixteen upper-middleclass pre-schoolers. In A1, A2 and A3 the token exchange was used; in B, only approval.

On the last day it was down to about 15 minutes—one-third the level of the end of the token period.

In A3, the token exchange was re-instituted. In only three days, attention snapped back from an average of 15 minutes to 45 minutes. However, by the end of A3, the students paid attention an average of 50 of the available 60 minutes.

A comparison of the record of these normals with the record of the hyper-aggressive boys is interesting. The increase in attention brought by the token exchange, from about 15 minutes to 50, is approximately threefold for the normal children; but for the hyperaggressive boys—who are disobedient and easily distracted—it is about eleven-fold, from 8 percent to 93 per-

cent of the time. The increase was not only greater, but the absolute level achieved was higher. This indicates strongly, therefore, that the more problematic the child, the greater may be the effect of token exchange on his behavior.

The high rates of attention were not due to the fact that each teacher had fewer children to work with. Individualized lessons were not enough. Without the token exchange, even three teachers could not hold the interest of 16 children 2 to 4 years old—at least not in reading, writing, and arithmetic.

Praise and approval were not enough as rewards. The teachers, throughout the experiment, used praise and approval to encourage at-

tention; they patted heads and said things like "Good," "You're doing fine," and "Keep it up"; yet, in B, when the token exchange was removed, this attention nevertheless ultimately declined by two-thirds. Social approval is important, but not nearly so powerful as material reinforcers.

Finally, it is obvious that if the reinforcers (movies, snacks, toys, or whatever) do not seem directly connected to the work, they will not sustain a high level of study. To be effective with young children, rewards must occur in a structured exchange in which they are given promptly as recompense and thus are directly connected to the work performed.

The Very Young Child

According to accepted educational theory, a child must be about six and a half before he can comfortably learn to read. But is this really true, or is it merely a convenience for the traditional educational system? After all, by the time a normal child is two and a half he has learned a foreign language—the one spoken by his parents and family; and he has learned it without special instruction or coaching. He has even developed a feel for the rules of grammar, which, by and large, he uses correctly. It is a rare college student who becomes fluent in a foreign language with only two and a half years of formal training—and our college students are supposed to be the brightest of our breed. Paul Goodman has suggested that if children learn to *speak* by the same methods that they learn to *read*, there might well be as many non-speakers now as illiterates.

What if the problem is really one of motivation? If we structured an exchange that rewarded them, in ways they could appreciate, for learning to

read, couldn't they learn as readily as 5-year-olds?

We decided that for beginners, the number of words a child can read is the best test of reading ability. In an experiment designed by Hamblin, Carol Pfeiffer, Dennis Shea, and June Hamblin, and administered at our Washington University pre-school, the token-exchange system was used to reward children for the number of words each learned. The results are given in Figure 5. Note that the 2-year-olds did about as well as the 5-year-olds; their sight vocabularies were almost as large.

There was an interesting side effect: at the end of the school year, all but one of these children tested at the "genius" level. On Stanford-Binet individual tests, their I.Q. scores increased as much as 36 points. It was impossible to compute an average gain only because three of the children "topped out"—made something in excess of 149, the maximum score possible.

In general, the lower the measured I.Q. at the start, the greater the gain—apparently as a result of the educational experience.

The Non-Verbal Child

What happens when ghetto children are introduced into a token-exchange system? At our Mullanphy Street pre-school, 22 Afro-American children—age 3 to 5—attend regularly. All live in or near the notorious Pruitt-Igoe Housing Project, and most come from broken homes. When the school began, the teachers were unenthusiastic about a token exchange, so we let them proceed as they wished. The result was pandemonium. About half of the children chased one another around the room, engaged in violent

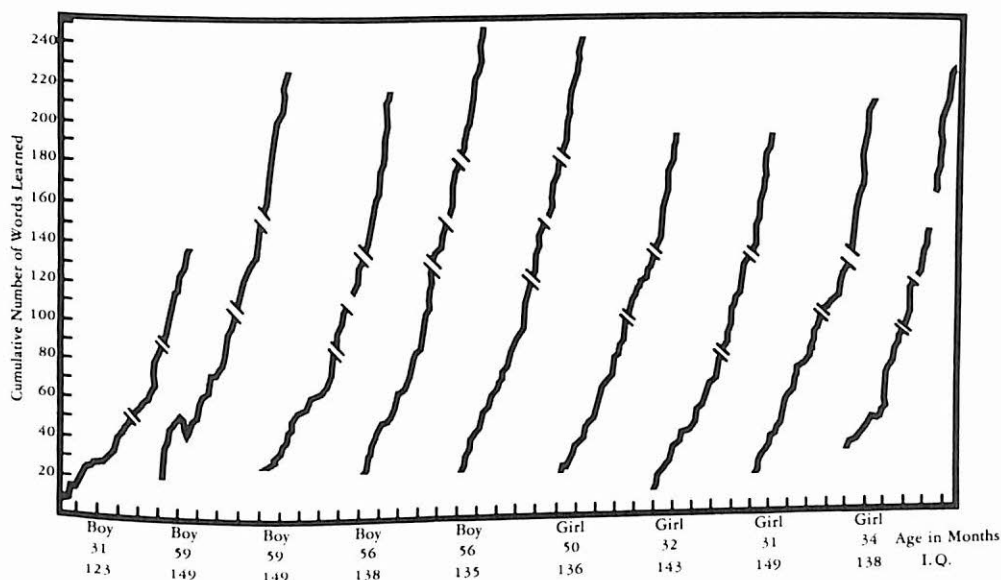


Figure 5

Number of sight-words learned through time by five 4- and 5-year-olds, and four 2- and 3-year-olds. Note that the younger children did about as well as the older ones—except for one boy whose I.Q. was somewhat lower than the others in the group. (Gaps indicate absences.)

arguments, and fought. The others withdrew; some would not even communicate.

After the third day, the teachers asked for help. As in the other experimental schools, we (Buchholdt and Hamblin) instructed them to ignore aggressive-disruptive behavior and to reward attention and cooperation with social approval and the plastic tokens, later to be exchanged for such things as milk, cookies, admission to the movies, and toys. The children quickly caught on, the disruptions diminished, and cooperation increased. Within three weeks of such consistent treatment, most of the children took part in the lessons, and disruptive behavior had become only an occasional problem. All of this, remember, without punishment.

Our attention was then focused upon the children with verbal problems. These children seldom started conversations with teachers or other students, but they would sometimes answer questions with a word or perhaps two. This pattern may be unusual in the middle classes, but is quite common among ghetto children. Our research has shown that children so afflicted are usually uneducable.

As we investigated, we became convinced that their problem was not that they were unable to talk as much as that they were too shy to talk to strangers—that is, to non-family. In their homes we overheard most of them talking brokenly, but in sentences. Consequently, we set up a token exchange for them designed specifically to develop a pattern of talk-

ing with outsiders, especially teachers and school children.

As it happened, we were able to complete the experiment with only four children (see Figure 6). During A1, the baseline period (before the tokens were used), the four children spoke only in about 8 percent of the 15-second sampling periods. In B1, the teachers gave social approval and tokens *only* for speaking; non-verbalisms, like pointing or headshaking, would not be recognized or reinforced. Note the increase in verbalization, leveling out at approximately 48 percent.

In A2 we reversed the conditions by using a teacher new to the school. The rate of talking dropped off immediately, then increased unevenly until it occurred in about 23 percent of the sample periods.

In B2 the new teacher reintroduced the token exchange for talking, and once more there was a dramatic rise: The speaking increased much more rapidly than the first time, ending up at about 60 percent. (This more rapid increase is known as the Contrast Effect. It occurs in part, perhaps, because the children value the token exchange more after it has been taken away.)

In the final test, we again took out the token exchange, and introduced yet another new teacher. This time the drop was small, to 47 percent.

We followed the children for three months after the end of the experiment. Their speech level remained at 48 percent, with little dropoff. This compares with the 40 percent talking rate for our other ghetto children, and the 42 percent rate for upper-middle-

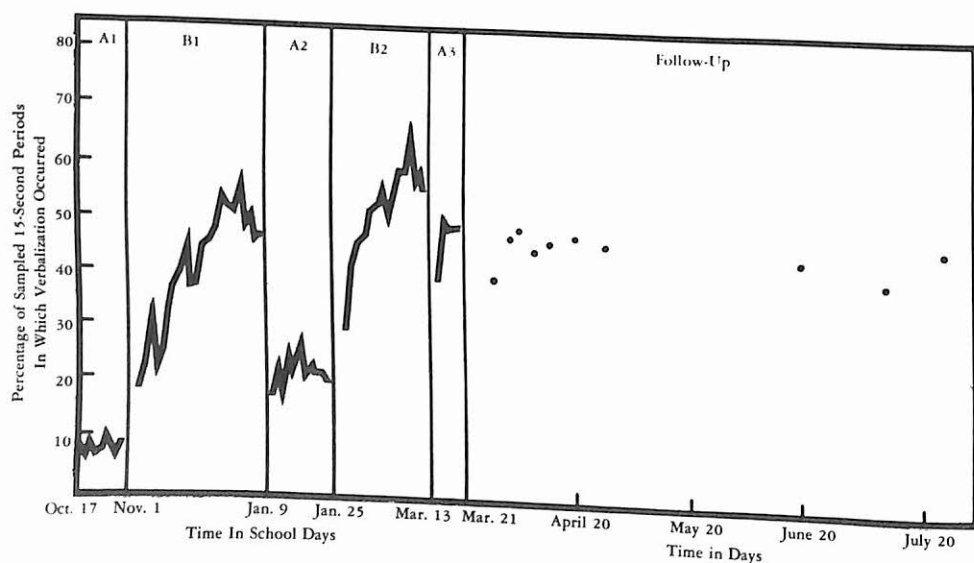


Figure 6

Percentage of sampled periods in which talking occurred through time for four non-verbal "culturally deprived" children, and through five experimental conditions. In each of the A conditions a new teacher was introduced, and she structured a token exchange for participation in lessons. In the B conditions, the teacher then structured a token exchange for talking. The follow-up was similar to the A conditions.

class children at the Washington University pre-school.

Frequency of speech, however, was not the only important finding. At the end of B1, the children spoke more often but still in a hesitant and broken way. By the end of B2, they spoke in sentences, used better syntax, and frequently started conversations.

Mothers, teachers, and neighbors all reported that the children were much more friendly and assertive. But some claimed that the children now talked too much! This could reflect individual bias; but there was little doubt that at least one child, Ben, had become an almost compulsive talker. He was given to saying hello to everyone he met and shaking their hands. So we terminated the experiment—what would have happened to Ben had we started *another* exchange?

This experiment shows that token exchange can bring on permanent behavior change, but that the culture must reinforce the new behavior. Talking is important in our culture, and so is reading; therefore they are reinforced. But other subjects—such as mathematics beyond simple arithmetic—are not for most people. For behavior to change permanently it must be reinforced at least intermittently.

Autism

The problems of autistic children usually dwarf those of all other children. To the casual observer, autistic children never sustain eye contact with others but appear to be self-contained—sealed off in a world of their own. The most severe cases never learn how to talk, though they sometimes echo or parrot. They remain dependent upon Mother and become more and more demanding. They develop increasingly destructive and

bizarre behavior problems. Finally, between 5 and 10 years old, autistic children ordinarily become unbearable to their families and at that point they are almost invariably institutionalized. Until recently, at least, this meant a rear ward to vegetate in until they died.

The breakthrough in therapy from autism came in 1964 when Dr. Ivar Lovaas and Dr. Montrose Wolfe and a graduate student, now Dr. Todd Risley, simultaneously developed therapy systems using well-established principles of operant conditioning. They were particularly successful with those children who randomly echoed or imitated words or sentences (this is called echolalia).

The therapy systems we have designed, developed, and tested, though similar in some ways to those developed by Lovaas, Wolfe and Risley, are quite different in others. First, we do not use punishment, or other negative stimuli. We simply terminate exchanges that reinforce the autistic patterns, and set up exchanges that reinforce normal patterns. Second, our children are not institutionalized; they live at home, and are brought to the laboratory for twenty minutes to three hours of therapy per day. Third, as soon as possible—usually within months—we get the children into classrooms where a therapist works with four or five at a time. Fourth, we train the mother to be an assistant therapist—mostly in the home, but also in the laboratory. These changes were introduced for several reasons, but primarily in the hope of getting a better, more permanent cure for autism.

The Etiology of Autism

Is autism hereditary, as many believe? Our studies indicate that this

is not the important question. Many mental faculties, including I.Q., have some physiological base. But the real issue is how much physiologically based potential is socially realized, for good or bad. As far as we can tell, the exchanges that intensify autism get structured inadvertently, often by accident; but once started, a vicious cycle develops that relentlessly drives the child further into autism.

When autism starts, the mother often reacts by babying the child, trying to anticipate his every need before he signals. Thus normal communication is not reinforced, and the child never learns to work his environment properly. But even if he doesn't know how to get what he wants through talking, he does learn, through association, that his oversolicitous and anxious mother will always respond if he acts up violently or bizarrely enough. And she must, if only to punish. He thus learns to play "Get mother's attention"; and this soon develops into "Get mother exasperated, but stop just short of the point where she punishes and really hurts." Here is an example (observed by Ferritor in the first of a series of experiments by the Laboratory's staff, not reported here):

Larry is allowed to pick out his favorite book. His mother then attempts to read it to him, but he keeps turning the pages so she can't. He gets up and walks away from the table. The mother then yells at him to come back. He *smiles* (a sign of pleasure usually but not always accompanies reinforcement). Mother continues to talk to the child to try to get him back for the story. Finally, he comes over to the table and takes the book away from her. She lets him and goes back to the bookcase for another book. He then sits down and she begins to read. He

tries to get up, but his mother pulls him back. Again. Again. She holds him there. He gets away and starts walking around the room. Goes to the toy cabinet. Mother gets up to go over and take a toy away from him. He sits on the floor. The mother comes over and sits down by him. He gets up and goes over by the door and opens it and tries to run out. She tells him he has to stay. He *smiles*. She resumes reading. He gets up and starts walking around the table. She grabs him as he comes by. He *smiles*.

A clinical psychologist who had tested Larry did not diagnose him as autistic, but as an educable mental retardate with an I.Q. of perhaps 30. Yet he had gaze aversion and we suspected that Larry, like other autistics, was feigning inability as a way of getting what he wanted from his mother, and then from other adults. He began to respond to the attractive exchanges that we structured for him, and as we did, he began to tip his hand. For example, at one point when his mother was being trained to be an assistant therapist, the following incident occurred:

Mrs. G. told Larry that as soon as he strung some beads he could have gum from the gum machine that was across the room. For about 10 minutes he fumbled, he whined, all the time crying, saying, "I can't." Finally, he threw the beads at his mother. Eventually, the mother had the good sense to leave the room, saying, "As soon as you string those beads, you can have your gum." With his mother out of the room, according to our observers he sat right down and, in less than 30 seconds, filled a string with beads with no apparent trouble.

Just two weeks later, after the mother had been through our 10-day training

program, they again had a "story time."

The mother begins by immediately asking Larry questions about this book (the same book used a few weeks before). He responds to every question. She gives approval for every correct answer. Then she tries to get him to say, "That is a duck." He will not say it intelligibly, but wants to turn the page. Mother says, "As soon as you say 'duck,' you may turn the page." Larry says "Duck" and turns the page. He smiles.

After seven minutes, Larry is still sitting down. They have finished one book and are beginning a second.

Most autistic children play the game "Look at me, I'm stupid," or "Look at me, I'm bizarre." These are simply attention-getting games that most adults repeatedly reinforce. Man is not a simple machine; he learns, and as he develops his abilities, he develops stronger and stronger habits. Thus, once these inadvertent exchanges get established, the child becomes more and more dependent, more and more disruptive, more and more bizarre, more and more alienated from the positive exchanges that are structured in his environment. What is sad is that the parents and the others in the child's life sense that something is terribly wrong, but the more they do, the worse the situation becomes.

It seems to those of us who have been involved in these experiments from the beginning that the exchange

techniques and theories we have used have without question demonstrated their effectiveness in treating and educating problem children. Watching these children as they go peacefully and productively about their lessons toward the end of each experimental series is both an exhilarating and humbling experience. It is almost impossible to believe that so many had been written off as "uneducable" by professionals, that without this therapy and training—or something similar—most would have had dark and hopeless futures.

But it is not inevitable that so many hyperaggressive or environmentally retarded ghetto children become dropouts or delinquents; it is not inevitable that so many autistic children, saddest of all, must vegetate and die mutely in the back wards of mental hospitals.

FURTHER READING SUGGESTED BY THE AUTHORS:

- The Analysis of Human Operant Behavior* by Ellen P. Reese (Dubuque, Iowa: William C. Brown Company, 1966).
- The Emotionally Disturbed Child in the Classroom* by Frank Hewett (Boston: Allyn and Bacon, Inc., 1968).
- Early Childhood Autism* edited by J. K. Wing (London: Pergamon Press, Ltd., 1966).
- Case Studies in Behavior Modification* by Leonard P. Ullman and Leonard Krasner (New York: Holt, Rinehart and Winston, Inc., 1965).

INTRODUCTION

This selection, like the previous one by Hamblin *et al.*, discusses the use of behavior modification techniques; but whereas the authors of the previous article concerned themselves primarily with education and classroom behavior, Goldiamond looks at a variety of problems. As you read, keep in mind the fundamental assumption of behavior modification: that human behavior is importantly and largely determined by the consequences of previous behavior. That is, people tend to repeat responses for which they have been rewarded or reinforced and they tend not to repeat other responses. The opponents of the behavior modification movement express a violent dissatisfaction with the very simple notion that reinforcement is a fundamental determinant of human behavior.

One of the basic concepts which Goldiamond mentions is that of *stimulus control*. A stimulus is said to have a *stimulus control* over a response if in some basic way that stimulus sets the occasion for that response, i.e., if the response generally occurs when the stimulus is present and doesn't occur when the stimulus is absent. At a very simple level the stimulus "How are you?" is generally followed in our society by the response "Fine" as long as the responder is healthy enough to hold himself erect. So "How are you?" could be said to have stimulus control over the response "Fine." Other examples of stimulus control are given in the article. Another fundamental concept used repeatedly is that of reinforcement, with which you are already familiar.

Note in the discussion on page 370 that punishment of bad handwriting had not been effective for improving the handwriting of the student who is being treated. Undoubtedly one reason why it was ineffective is that the alternative and desired response of legible handwriting simply was not available to the subject. He didn't know how to write legibly, so his parents and teachers could punish his bad writing as much as they wished without producing the desired change. They were unable to change the quality of the writing simply because the alternative response, writing well, was not in the subject's repertoire. Remember in the previous article that Hamblin *et al.* commented that simply ignoring aggres-

sion probably wouldn't have worked but that ignoring aggression while at the same time rewarding cooperative behavior did work.

Although Goldiamond's article discusses, rather informally, the various cases on which the author worked, it does give some idea about how you can break down complicated behavior patterns into simpler ones which in turn can be modified by simple contingencies of reinforcement.

SELF-CONTROL PROCEDURES IN PERSONAL BEHAVIOR PROBLEMS

Israel Goldiamond

Another advantage of direct treatment of disordered behavior lies in the breaking of possible symptom→underlying cause→symptom circularity. For example, Wickes, in his account of a behavioral treatment of enuresis, points out that often the "anxiety" considered the underlying cause of the problem behavior disappeared when the bed-wetting stopped. In many cases, the enuresis, by aggravating interpersonal, parent-child relationships, may have been the cause of the "anxiety" rather than the symptomatic result. Thus, at least in some cases, behavioral therapy is not only a convenient, but an ultimately effective treatment of the person's "problem" as a whole.

Perhaps the most refreshing aspect of these articles is their specificity in telling the rehabilitation worker, doctor, parent, teacher, or patient—whoever is concerned with the modification of disordered behavior—precisely what they are to do to alter the behavior in question. The contrast between such behavioral therapy and the more esoteric procedures involved in other forms of therapy provides a new and exciting approach to old and troublesome problems.

The present discussion will be concerned with the application of self-control procedures to the solution of certain limited behavioral problems.

Often one person comes for help from another because he cannot cope with problems that face him. The appropriate behaviors are not available.

Reprinted with permission of author and publisher (this is available as a separate from the journal): Goldiamond, I. "Self-Control Procedures in Personal Behavior Problems." Psychological Reports, 1965, 17, pp. 851-68.

The means by which the behavioral deficit can be overcome are varied. Simple instructions often suffice, as when the subject cannot study because he does not have the assignment. On the other hand, the subject may not be able to study because he cannot allocate his time appropriately, because he daydreams at his desk, or because he engages in other behaviors which come under the general heading of lack of self-control. In these cases, simple instructions will not remedy

the deficit since the subject himself knows what it is. He has often tried to instruct himself to behave appropriately, but with little success. Indeed, the numerous jokes surrounding New Year's resolutions indicate both the prevalence of the problem, and the ineffectiveness of its instructional solution, whether imposed by others or by one's self in self-instruction.

The specific behavioral deficit, or presenting problem, is often part of a larger context of deficits. Rather than trying to overcome the presenting deficit directly, the therapist may turn his attention to other, "deeper" behaviors or behavioral deficits. In this case, the presenting problem is considered a symptom, by exactly the same defining operations that make a skin rash a symptom. Here, the dermatologist states that to get rid of the rash he will treat in addition something else, possibly a blood imbalance, rather than only treating the rash itself (Goldiamond, Dyrud, and Miller, 1965). For example, we know of one case where a woman assumed a fetal posture for three days after an argument with her husband. She was restored to mobility by direct modification of this behavior. It can be argued that a woman who maintains control over her husband by such extreme behaviors is so lacking in more appropriate behaviors that her treatment will require considerably more than the two hours which restored her mobility. This case may be an example of treating a symptom, since behaviors other than the presenting complaint may also require modification. However, the "underlying" problem can still be considered as a behavioral one: in this case, the absence of those behaviors whereby wives normally control their husbands. Yet this general deficit also may be treated di-

rectly, as in our research on stuttering where we have, within the laboratory, replaced stuttering by fluent and rapid speech in 30 of 30 cases run thus far (Goldiamond, 1965b). Indeed, one of our stuttering patients, who had been suicidal, became able to read bedtime stories to her children at home, and certain other personal problems at home cleared up because her stuttering cleared up. Some of her other behaviors were accordingly symptomatic of stuttering (Goldiamond, 1965a), as we have defined "symptom."

If there is a danger in premature assignment of behavioral deficits as symptomatic, there is also a danger in premature assumption that the alteration of the presenting problem is the final solution. Further analysis in light of current knowledge will undoubtedly both modify and confirm practice in this area.

Classification of behavior as a problem for treatment or as a symptom may also be an economic or contractual matter. For example, in a marital problem, the presenting complaint may be cleared up in a short period of time, but other problems are sometimes uncovered which may require extensive treatment. At what point is the implicit contract between patient and therapist (cf. Sulzer, 1962) to treat the marital behaviors extended to behaviors in other areas? The answer to this question must depend on the extent to which the subject can afford the treatment or can afford not to get it, that is, can do without it. It would be nice to have a new car when the present one seems to require extensive repairs, but there may be other considerations such as a piano or a child's education. For going to and from work, minor adjustments may be sufficient. The economy may also be behavioral: is it worth the upset?

The present discussion will be limited to cases where the concern was with a specified behavioral problem. These cases should be interpreted in the context of the foregoing discussion: namely, that the procedures used are not intended to question other more extensive procedures, which may be necessary for other kinds of behavioral deficits.

The discussion will be concerned with self-control (Skinner, 1953) and procedures for its establishment. The procedures to be discussed center around the position that behavior is not an emergent property of an organism nor a property solely of its environment, but is described by a functional relation between the two. More technically, given a specified behavior B and a specified environmental variable x , a lawful relation can be found, such that $B=f(x)$, under certain empirical constraining conditions c . This implies that when the constraints c are set up, and x is set at a stipulated value, then B will have a stipulated value, given by the value of $B=f(x)$. When the experimenter sets x at that value, *he* will get the B stipulated. This defines the experimental control of behavior which has been demonstrated repeatedly in operant and other laboratories. When the subject himself sets x at that value, *he* will get his own B , as stipulated. This defines self-control.

If you want a specified behavior from yourself, set up the conditions which you know will control it. For example, if you cannot get up in the morning by firmly resolving to do so and telling yourself that you must, buy and set an alarm clock. Within this context, the Greek maxim, "Know thyself," translates into "Know thy behaviors, know thy environment, and know the functional relation between

the two." Although the relation between an alarm clock and waking up is a simple and familiar one, other relations are neither this simple nor this familiar. There have, however, been developed in laboratories of operant behavior a body of known functional relations between behavior, and programs and other procedures which can alter even more complex behavior systematically. Self-control derived from such research can take at least two forms. One is to instruct the subject to set up the procedures which change his environment and which thereby bring his behavior under different control. I shall present some cases to this effect. Another form is to train him in the functional analysis of behavior, and have him try to determine for himself the procedures which he should apply. This approach will also appear in the following cases.

Inherent in both types of self-control is the problem posed by the tremendous gap between theory and practice. The same theory may dictate numerous alternative methods or solutions, but all may not be equally available, practical, or applicable. The operant paradigm suggests that there are at least twelve different ways to maintain or attenuate behavior (cf. Holz and Azrin, 1963). Which are appropriate to the problem?

One way of selecting effective practical measures is to have S report back to E every week with his results. This hour becomes a session for analysis of data and discussion of changes in procedure. In the laboratory, operant procedures are so arranged that relations between ongoing behavior and its conditions are continuously observed and recorded. Through successes and failures, Es may learn to analyze behavior and conditions and may develop a "feel" for their data,

as do other behavioral practitioners in interaction with their subject matter, for example, skilled psychiatrists. Hopefully, such a program of systematic trial and analysis will sensitize *S* to his own behavior and his own conditions. By training *S* in control procedures to the extent that these exist and are applicable, we are providing for self-enhancement and self-actualization (Rogers, 1951). Of the individuals who can apply control procedures, *S* is the one most concerned with his behavior and is most in contact with it, its conditions, and its consequences. Initially *E* is the consultant, and eventually *S* becomes his own *E*. The procedures may be limited to *Ss* who are intellectually capable of such analysis or who are not otherwise incapacitated. Our *Ss* were mainly college students. Where systematic training in behavior analysis was used, the sessions started with individual tutorials in behavior analysis, homework assignments from standard texts (Holland and Skinner, 1961), and readings. Given this intellectual base, we could move on to discussions of the problem in question.

Our first cases were referrals from clinical psychologists who felt that we should work on some of the simpler overt problems, while they tackled their deeper meanings. One of these was a young man who was overweight, and another was a girl who had difficulty studying.

These two problems yielded, for these *Ss*, to procedures involving *stimulus control*. The program with the young lady started with human engineering of her desk. Since she felt sleepy when she studied, she was told to replace a 40-watt lamp with a good one and to turn her desk away from her bed. It was also decided that her desk was to control study behavior.

If she wished to write a letter, she should do so, but in the dining room; if she wished to read comic books, she should do so, but in the kitchen; if she wished to daydream, she should do so, but was to go to another room; at her desk she was to engage in her school work and her school work only.

This girl had previously had a course in behavioral analysis and said, "I know what you're up to. You want that desk to assume stimulus control over me. I'm not going to let any piece of wood run my life for me."

"On the contrary," I said, "you *want* that desk to run you. It is you who decides when to put yourself under the control of your desk. It is like having a sharpened knife in a drawer. You decide when to use it; but when you want it, it is ready."

After the first week of the regimen, she came to me and gleefully said, "I spent only ten minutes at my desk last week."

"Did you study there," I asked.

"Yes, I did," she said.

"Good," I said, "let's try to double that next week."

For the next few weeks we did not meet, but she subsequently reported that during the last month of the semester she was able to spend three hours a day at her desk for four weeks in a row, something she had been unable to do previously. When she sat at her desk she studied, and when she did other things she left her desk. The variable maintaining this increase in behavior as the semester drew to an end was apparently the forthcoming final examinations.

With regard to the young man who overate, stimulus control, chaining, and withdrawal of reinforcement were used. The stimulus for overeating is normally not food (Ferster, Nurnberger, and Levitt, 1962). In our cul-

ture, food is normally hidden; it is kept in a refrigerator or cupboard. In the cafeteria, where it is in the interests of the management to get people to eat, food is exposed.

The initial strategy for slimming the young man was to bring his eating behavior under the control of food alone, since food is normally not available as a stimulus. He was instructed to eat to his heart's content and not to repress the desire. He was, however, to treat food with the dignity it deserved. Rather than eating while he watched television or while he studied, he was to devote himself to eating when he ate. If he wished to eat a sandwich, he was to put it on a plate and sit down and devote himself exclusively to it. Thus, reinforcing consequences such as watching television or reading would be withdrawn when he engaged in the behaviors of preparing the food, eating, and cleaning up. Responding to the refrigerator in between meals resulted in withdrawal of such consequences, as did going to the refrigerator while watching television. Television, studying, and other stimuli would lose their control of initiating the chain of behaviors and conditions that terminated in eating. Within one week, the young man cut out all eating between meals. "You've taken the fun out of it," he said to me. We then worked on the contents of the meals as well, and he stopped attending sessions. I met him about three months later: he was considerably slimmer and remarked that he needed neither me nor the clinical psychologist to solve his problems. He could handle them himself.

No claim is made that all problems should be treated in this manner, or that the Ss had no other problems. The aim was to alter the specified

behavior. We started out with the simplest procedures. Had these not been effective, we would have tried others. Some more complex cases will be presented next.

An interesting aspect of these and other cases as well was the fact that in a very short time the Ss ran off by themselves to apply the procedures they had learned. In some cases, I would have preferred more extensive interchange, and wondered how clinical psychologists were able to keep Ss coming week after week. Finally I attributed the tenure of the relationship to what might be called the Scheherazade effect. Scheherazade, as you will recall, became the consort of a king who killed each bedmate after one night, having generalized the infidelity of a previous wife to all women. Scheherazade told him a story on the first night, which was not completed by dawn. The king paroled her for the second night to hear the rest of the story, and having been reinforced, she repeated her behavior. The schedule maintained such behavior for 1001 nights, and the result is known as the *Arabian Nights*.

Few things are more interesting and will sustain behavior better than support for talking about one's self; one is never finished in 50 minutes. Hence, such discussions may maintain therapy sessions and allow the therapist to interact with the patient over an extended period of time. An individual tutorial may serve the same function.

Marital Case

The husband in this case was a young man, 29, who was working on his master's degree. His wife was taking my course in behavioral analysis, and they both decided that he

should come to see me about their marriage, which both wanted to maintain. The issue, as *S* told me, was that his wife had committed the "ultimate betrayal" two years ago with *S*'s best friend. Even worse, it was *S* who had suggested that the friend keep his wife company while he was in the library at night. Since that time, whenever he saw his wife, *S* screamed at her for hours on end, or else was ashamed of himself for having done so and spent hours sulking and brooding. Since the events that led to the "betrayal" were an occasion for bringing home the first lesson on the consequences of behavior, we started from there.

RELATION OF BEHAVIOR TO ITS CONSEQUENCES

Early discussions concerned the analysis of behavior in terms of its consequences. *S*'s behavior provided stimuli for his wife's behavior. If he wished his wife to behave differently to him, then he should provide other stimuli than the ones which produced the behaviors he did not like. There was considerable analysis of such interactions. This conceptualization of behavior was apparently new to *S*, who took detailed notes; and I have discovered it to be new to many other *S*'s as well.

STIMULUS CHANGE

Altering the consequences of operant behavior will alter the behavior itself. However, this process may take a considerable amount of time. One of the most rapid ways to change behavior is by altering the conditions under which it usually occurs. This is called *stimulus change*, or the effects of novel stimuli. If the novel stimuli are then combined with new behavioral contingencies designed to

produce different behavior, these contingencies are apt to generate the new behavior much more rapidly than they would in the presence of the old stimuli.

As part of the program of establishing new stimuli, *S* was instructed to rearrange the use of rooms and furniture in his house to make it appear considerably different. His wife went one step further and took the occasion to buy herself a new outfit.

ESTABLISHMENT OF NEW BEHAVIOR

Since it was impossible for *S* to converse in a civilized manner with his wife, we discussed a program of going to one evening spot on Monday, another on Tuesday, and another on Wednesday.

"Oh," he said, "you want us to be together. We'll go bowling on Thursday."

"On the contrary," I said, "I am interested in your subjecting yourself to an environment where civilized chit-chat is maintained. Such is not the case at a bowling alley."

I also asked if there were any topic of conversation which once started would maintain itself. He commented on his mother-in-law's crazy ideas about farming. He was then given an index card and instructed to write "farm" on it and to attach a \$20 bill to that card. The \$20 was to be used to pay the waitress on Thursday, at which point he was to start the "farm" discussion which hopefully would continue into the taxi and home.

STIMULUS CONTROL

Since in the absence of yelling at his wife *S* sulked, and since the program was designed to reduce yelling, *S*'s sulking was in danger of increasing. *S* was instructed to sulk to his

heart's content, but to do so in a specified place. Whenever he felt like sulking, he was to go into the garage, sit on a special sulking stool, and sulk and mutter over the indignities of life for as long as he wished. When he was through with his sulking, he could leave the garage and join his wife. He was instructed to keep a daily record of such behavior and bring it to each session. The graph is presented in Figure 1. Sulking time had been re-

markedly altered the perceptual configuration of the room.

RECORDS

Daily notes of events were kept in a notebook, as was the graph. *S* took notes of the discussions with *E*. These notes were discussed at each weekly session.

One of the notions which *S* held very strongly was that his wife's behavior stemmed from some inaccessible source within her, and that many of his own behaviors likewise poured out from himself. In this context, the final sharp rise in the sulking curve was discussed. "The whole procedure won't work," he said, "my wife doesn't need me as much as I need her." The psychiatric message was that he had no control over his wife, but I chose to ignore this message in favor of a didactic one on the behavioral definition of needs. He was asked how he knew what his wife's needs were. Was he an amoeba slithering into her tissues and observing tissue needs? Was he a mind reader? After my repeated rejection of subjective definitions of needs, he redefined the problem behaviorally, namely, that his wife behaved a certain way less than he did. He said that stated this way it sounded silly, but I said, "No, it's a problem to you and not silly."

What were these behaviors? They apparently included such dependency behaviors as asking him to do things for her. "When was the last time she asked you to do something for her?" I asked. He replied that the previous day she asked him to replace a light bulb in the kitchen. Had he done so, I asked. No, he said. He then was asked to consider the extinction of pigeon behavior and took notes to the effect that, if he wished his wife to act

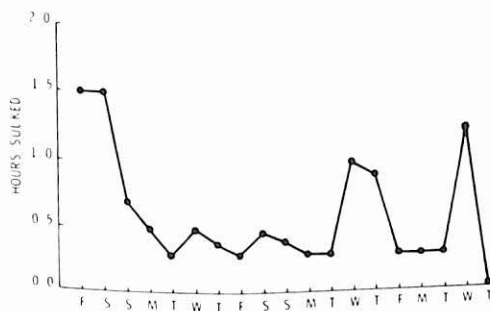


Figure 1

Graph kept of sulking behavior.

ported as 7 hours on the preceding day, and, with occasional lapses, it was reported as dropping to less than 30 minutes before disappearing entirely. The reported reversals and drops were occasions for discussions.

Since the bedroom had been the scene of both bickering and occasional lapses, the problem was presented of changing its stimulus value when conjugality was involved. If this could be done consistently, eventually the special stimuli might come to control such behavior. The problem was to find a stimulus which could alter the room entirely and would be easy to apply and withdraw. Finally, a yellow night light was put in, was turned on when both felt amorous, and was kept turned off otherwise. This light

helpless, he should reinforce dependency by doing what she asked.

A discussion on needs and personality ensued. "If by personality all that is meant is my behavior," he said, "then my personality changes from one moment to the next, because my behavior changes," he stated.

"I should hope so," I said.

"Well, what is my true personality; what is the true me?" he asked.

"Do you have a true behavior?" I asked.

He reported this as a viewpoint he had never considered; his previous training had been in terms of being consistent to his self, and of searching for "thine own self (to which he could) be true." He took extensive notes.

The next week he came in and stated: "I did something last week that I have never done before in my life. When I teach in classrooms I am able to manage my students, but when I talk to tradespeople I find I am very timid and allow myself to be cheated. Well, last week my carburetor gave out. I knew if I went to the garage they would make me buy a new one even though I have a one-year's guarantee. I sent my wife down to the garage instead. She is a real scrapper. She came back with a new carburetor. It didn't cost us a cent. Why should I have to be all things to all men? In school I control things, but with tradespeople I don't. So what?"

These weekly sessions continued during ten weeks of the summer term. After the initial training, S was assigned homework along with his wife who was taking the course in behavioral analysis. The weekly discussions were centered around behavioral analysis and how it might apply to his problems.

During the course of one of the

sessions, S started to talk about his childhood and was summarily cut-off.

"Shouldn't I talk about this with a psychologist?" he asked. "Isn't this one of the things that interests you? Doesn't it affect me now?"

"Look," I said, "a bridge with a load limit of three tons opens in 1903. The next day, a farmer drives eighteen tons over it; it cracks. The bridge collapses in 1963. What caused the collapse?"

"The farmer in 1903," he said.

"Wrong," I said, "the bridge collapses in 1963 because of the cracks that day. Had they been filled in the preceding day, it would not have collapsed. Let's discuss the cracks in your marriage."

At the end of the period, there was no sulking in the garage and the partners were able to commune.

Study Case 1: Handwriting

S was a seventeen-year-old high school senior of normal intelligence, who was 28th in a class of 28 and whose handwriting was illegible. He was a referral from a school psychologist.

An example of his early handwriting is given in Figure 2. Lettered lines of translation alternate with curvilinear lines of handwriting. During the first session, I asked S to sit at my desk and write from dictation. He leaned forward to write, but no part of his hand or arm touched paper or desk; the entire force of his shoulder and arm was transmitted to the pencil point, making fine control impossible. Since the primary grades, no one had ever observed his writing behavior; they had, however, criticized its product. I instructed him to keep his arm on the table and to manipulate his pencil from the wrist and fingers.

(Tough Draft)

English-IE Theme
Theme
Theme

This person is of average height, a little
~~This person is of average height, a little~~
 too fat in many ways. He is somewhat
 too loud and always using his ~~strong~~ tongue. It
 never ceases to continue always
 never ceases to continue always
 friendly but obnoxious. He talks ~~touchy~~ with
 a somewhat southern twang in his voice,
 which I imagine is put on. In many ways,
 which I imagine is put on. In many ways,
 he sums up this class; a group of
 immature people, which need to grow
 up and learn to be Americans first,
 and above all, learn to keep their
 mouths shut. I ~~touch~~ (only) this unbearable (creature?), for
 mouse's sake. I ~~touch~~ this unbearable for
 he is the most obnoxious of all.
 he is the most obnoxious of all.

Figure 2

Sample of initial handwriting; translation appears above line.

Some simple physics was explained to him. Sheets of onion skin were interlaced with carbon paper, and he was given exercises requiring modulation of force so that he would go through five sheets, four sheets, three sheets, and two sheets. He was instructed to print.

At a later session, his letter size being erratic, he was asked to letter the familiar, "Now is the time . . ." phrase. Figure 3 presents that initial attempt. The paper is lined, and the writing starts out filling up the space

between the lines. The writing becomes smaller and smaller. This tendency can be traced to the letter T. In all cases, T is the same size as the preceding letters, and is also a capital. Since capital letters are followed by small letters, these T's control the size of the small letters that follow them. The first T in the second line produces a row of smaller letters following it, and the first T in the third line takes off from this size and again cuts down the size of the following letters until the final T produces a

now IS The Time ALL
good men to come to
The ade of their party.

Figure 3

Stimulus control of ensuing letter size by writing T as capital of same size as preceding letter.

tinier Y. S was instructed to write his T's so that they were larger than the preceding letters. Figure 4 reiterates the control this letter had over the following letters. The letters which follow T are smaller than T, but since T is above the line they stay within the line. The effect has been reversed, demonstrating a causal relation. Incidentally, since T is the second most frequently used letter in English, it is a powerful source of control. In a later session, S was instructed to differentiate between capital and small T.

The passage presented is the same as that in Figure 2. The change in legibility is evident. The numerous spelling errors are of interest. As long as the writing is undecipherable, spelling errors can not be noted and corrected.

Many letters contain similar forms. For example, the letters a, b, d, o, p, and q all contain a complete circle of the same size. These circles are modified in letters such as c and e. S was instructed to bring an assortment of buttons to the session and a suitable button was found for him to use for these letters, with other buttons for other letters. He was also instructed in other principles of writing. He practiced at home and brought his material in weekly. After a period of lettering, S was instructed to link his letters in an effort to produce cursive writing.

This person is
OF average health
a little over
size in many
ways He is
some what To
loud and alw-
ays Talking
And never ceases
To Talk He Talk

Figure 4

Reversal of effect by writing T above line.

His handwriting improved markedly during this period, and he rose from 28th in his class to 13th. The undergraduate assistant² who worked with him at home paid him money for lines completed and, as long as this procedure was used, S's handwriting was legible and showed evidence of training. When this pay was dropped, the handwriting deteriorated. Although legible handwriting was now contingent upon reinforcement, differential reinforcement alone would not have produced the new behavior. Indeed, without making them contingent upon a program to alter behavior, differential consequences may be ineffective and may result in aversive control, as did the nagging and the poor grades he consistently obtained. Now that S has the new be-

² I wish to express my appreciation to Mr. Richard David, then an undergraduate psychology major, who brought S to and from sessions and assisted in them.

havior, differential consequences can be applied to maintain it; hopefully, these will be provided by society. The requirement of extrinsic consequences was probably related to difficult conditions at home. He had no desk or work place, and he lived alone with his mother.

Several conclusions can be drawn from this case. One is that observation of the behavior itself may on occasion be far more useful than observation of its end product. Another is that behavior may be controlled by the very stimuli that the behavior itself produces, as in the size of the letters following T. A third is that merely establishing a behavior will not necessarily maintain it. The consequences which maintain it must be considered, but, as a fourth conclusion, these consequences should often be related to a program of behavioral modification. Penalizing his poor writing behavior, as his teachers had done, did not eliminate that behavior. It did, however, bring him in for treatment.

Study Case 2: Study Problem

This case involved a junior in college who was being expelled because of his academic record. Inasmuch as I had given him one of his D's, he came to me for advice. I told him that I would try to get him reinstated, providing he put himself under my control. He agreed and was given a conditional year. S's parents were professionals. Two siblings were at prominent Eastern universities. S was alert, and his I.Q. tests were within the range of college populations.

S had never actually studied. Accordingly, a self-control regimen for studying was introduced following some of the procedures which were discussed earlier. A daily record was

kept of the total number of hours spent in study, for each course and for the total day.

At that time S studied for an average of six hours a day, but his study hours were not evenly distributed among the various classes. Figure 5 presents the

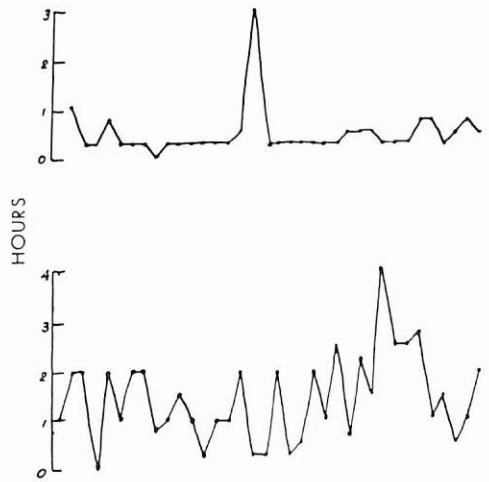


Figure 5

Daily study charts for class with one examination (upper), and class with daily recitations (lower).

study curves for two different courses. The upper curve depicts minimal studying, except for one peak. Guess when the test was given? This fixed interval curve characterizes the behavior of pigeons, as well. The lower part of the curve represents the studying pattern for a language class which involved a daily recitation. S studied at regular intervals. These curves so impressed me that I instituted daily quizzes in my classes, and was gratified to see the averages rise.

Every week, S brought his notebooks, his examinations, and his weekly records for inspection and discussion. Different procedures for keep-

ing notes were developed for each course. In the foreign language, for example, a separate sheet was kept for masculine and feminine nouns, and for each verb class. Grammatical classes represent similar ways of treating a word. By putting words of the same grammatical class together, there may be generalization from the changes learned for one word to the changes learned for another. Flash cards may not lend themselves to such generalization.

S was taught to outline social science texts. He was asked to be a detective and resurrect the author's outline (where he had one). Red and black ink were used for headings and entries. The outlines were topical rather than sentence outlines. This required restating the sentences in his words, rather than copying them from the book.

In English, we were both at a loss. Several novels were required, and neither one of us knew what should be considered. We used the conventional procedure to find out, namely, the first exam. This told us what the instructor considered to be the terminal behaviors.

At the end of the first semester, S's grades changed from the two D's and three E's of the preceding semester to two C's and two D's. Although this was an improvement, one class was dropped. I was disturbed since I felt that, with this much study and careful outlining, S should have obtained A's. A more detailed analysis was then undertaken.

S was taking a course in international trade, for which a knowledge of geography is requisite. He said that he knew his geography quite well and was asked to draw a map of South America. Figure 6 presents the map he drew. Brazil extends from sea to

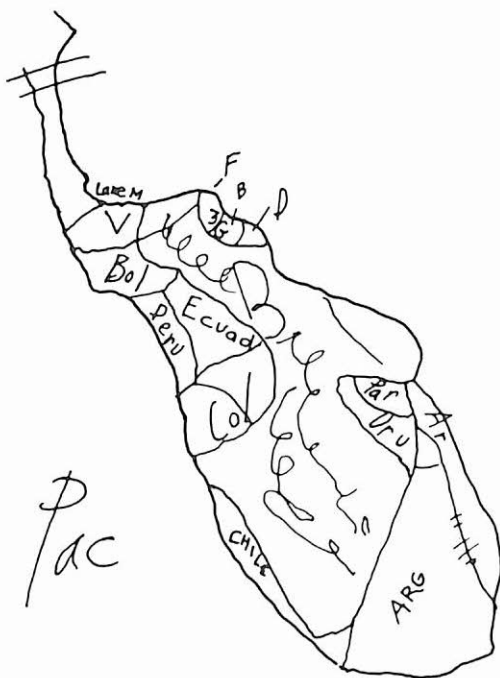


Figure 6

Map of South America drawn from memory by college junior taking a course in international trade.

shining sea. The body of water above Venezuela is designated Lake Maracaibo, and Bolívia is north of Peru and abuts on Venezuela.

I asked S what grade he would assign himself for this map, and he looked at it confidently and said, "Oh, 75 per cent or a C." I said that I would give him 20 per cent. Incidentally, in work with other S's since then who have trouble with history, I have often found that they lack knowledge of geography, and even of map-reading. Accordingly, historical movements become disconnected facts which have to be learned for each case.

This map may be used to exemplify a behavioral definition of stupidity.

Many behaviors require other behaviors as prerequisites for their acquisition and maintenance. It is possible that, in one form of stupidity, the prerequisites to the attempted behavior are absent, as well as those discriminative behaviors which differentiate the presence of adequate behaviors from their absence. Stated otherwise, the person we call stupid is lacking certain behaviors, but behaves as if he is not so lacking. He does not know to ask. He confidently undertakes assignments and often starts out successfully. However, where the new behaviors require older ones and these are missing, we may obtain the egregious blunders we call stupid. The blunders may be unpredictable to us simply because, in a long sequence, there are too many places in which deficits can occur for us to have come across every one.

These behavioral deficiencies were found in one area after another. Since *S* was taking courses which had as prerequisites other courses, which he had passed with D's, he was being required to acquire new behaviors when the prerequisite foundation for them was rather shaky. Accordingly, we "regressed" to the freshman texts in those courses in which he had obtained a D.

But there were deficiencies here, as well. In the economics texts, Humpty-Dumpty was quoted to the effect that words meant what *he* intended them to mean. The point here was that the economist's use of words might differ from their common usage. Humpty-Dumpty appeared in red in *S*'s notes. *S* knew that Humpty-Dumpty was an egg who fell. I asked why the egg led his paragraph, and could obtain no answer. It turned out that *S* had not read *Alice in Wonderland*, nor any of the childhood classics, nor for that matter *Tom Swift*, nor even comic

books. He simply did no home reading as a child.

I had obtained excellent control over *S*'s behavior, but this was like successfully getting someone to work six hours a day copying Chinese letters with a brush, without his ever having learned how to hold a brush or what the significant calligraphic nuances were. I suggested a program of visiting the art galleries, listening to concerts, reading the classics, and otherwise acquiring the behaviors relevant to our cultural heritage.

There are two types of behavioral sequences. In one type of sequence, called the *chaining* sequence, a chain of behavior is maintained by the consequences attached to the last element in the chain. Thus, Lundin (1961) reports a rat who went through various behavioral gyrations, then up five stories and down an elevator. All of these behaviors were maintained by the food he received at the end. In this type of sequence, the order of training is the *reverse* of the chronological order in which the sequence of behavior is performed. The pressing of the lever for food was established first. This was made contingent upon manipulation of the elevator. Then the ride in the elevator was made contingent upon the preceding step and so on. The entire chain was maintained by the food. Thus, if the product of education is not reinforcing, the behaviors which lead to it may not be maintained. Chaining was also exemplified in the weight reduction case, when watching television no longer served as a consequence of going to open the refrigerator.

Another type of sequence, which we shall call the *systematic* sequence, is exemplified in the case of this student. A systematic sequence can be compared to a course or an educa-

tional curriculum: the acquisition of one behavior depends upon the *prior* existence of another, just as the acquisition of new knowledge depends upon a grasp of other presupposed knowledge. Thus, in order to learn algebra, we must know how to read. In order to learn to read, a certain degree of socialization must first occur. Such curricula exist not only in academic subjects, but are implicit in other types of behavior as well. It is entirely possible, for example, that certain behaviors necessary for marital success presuppose the existence of other interpersonal behaviors, whose acquisition depends upon the existence of yet other behaviors.

Stated otherwise, there are behavioral curricula involved in almost all behaviors. Where the present behavioral deficit exists because an earlier behavior was not acquired, a procedure which attempts to correct the ongoing deficit must consider some of the earlier deficits of which the present one may be an outgrowth. When this is true, in contrast to the first marital case presented, discussion of childhood may be necessary.

The cases presented here have involved different behavioral deficits. Although the content or the topography of these and other behaviors differ, the functional relations of such differing behaviors to their environment may be similar. This may be true not only where the differences are in such categories as marital or academic, both of which involve human behavior, but also where the differences categorize species. Pigeons peck, and people talk. Topographically these are different behaviors. However, if the consequences which maintain pecking are scheduled in a certain manner, and the (quite different) consequences which maintain

speech are scheduled in the same manner, then the differing behaviors of pecking and speech will undergo similar changes in rate. It is upon this functional, or dynamic, similarity in the relation of behavior to its environment that the possibility of the extension of procedures from the laboratory to the clinic rests. It also suggests that problems in the clinic may be used for research in the laboratory.

The cases presented here demonstrate a simple application of certain self-control procedures derived from the laboratory. As was explicitly indicated earlier, they are not intended to supplant or question other more complex procedures (for a more extended analysis see Goldiamond, Dyrud, and Miller, 1965). However, we are currently examining some of these complex procedures, and are discovering that the explicit language of the laboratory may be very useful in analyzing and describing some of the behavioral transactions and changes that go on in other forms of psychotherapy (Goldiamond, Dyrud, and Miller, 1965). Developments in other areas where explicit analysis is utilized may be considered for their relevance to psychotherapy, and psychotherapy, by a reverse lend-lease, may suggest areas of investigation under more controlled procedures (Goldiamond, 1966).

Laboratory research has necessarily been characterized by a simplicity of procedures and concepts, and their extension to the solution of complex human problems requires considerable precaution and careful examination. Nevertheless, these procedures and concepts may provide methods for the analysis and restatement of complex problems in observable and manipulable terms and may thereby assist in the explicit assessment of behavioral change and effectiveness.

The cases presented here involve behavioral problems which could be analyzed by the *Ss* themselves. This training of *S* to become his own therapist is one of the goals of most branches of psychotherapy. The method used to accomplish this will depend on the state of the art, the nature of the problem, *S*'s past history, and social and other constraints upon *S*'s behavior. In some cases, these factors may dictate a strategy of not instructing *S*, or not having him de-

fine the problem or discover its solution immediately, since such a procedure may disrupt other behaviors and the consequences currently maintaining them. The course of treatment might then be considerably different from any of those discussed here. Other problems and possible procedures could be cited as well. But the cases presented here suggest that, in some areas at least, simple procedures can lead to complex changes.

INTRODUCTION

B. F. Skinner is in many ways the intellectual father of the behavior modification techniques described in the previous two articles. The following excerpt, taken from a book by Skinner published in 1953, was written before the more detailed studies which you have just read had been completed. Although the chapter appears a bit naive in places, it does an honorable job presenting the extreme point of view that *all* of human behavior can properly be thought of as responses with strengths which are changed as a result of reinforcement or non-reinforcement.

In this chapter Skinner considers a variety of techniques available by which someone can influence future consequences of his own behavior. He cites the alarm clock as a simple example of this sort. One of your editors found a simple alarm clock ineffective for controlling his morning behavior and had to resort to a more elaborate stratagem: he locked the alarm clock, together with a pot of coffee, inside a wooden chest which was then closed with a combination padlock. When the alarm clock sounded he had to turn on a light and get sufficiently awake to open the combination padlock. As soon as he had done so, and turned off the alarm clock, the coffee pot was immediately available to help sustain the waking state. Such extreme arrangements may be rather unusual but they indicate the length to which one may go in setting up special environments in which his own behavior will be differentially controlled.

As you read this article think of additional examples from your own experience to add to the lists given by Skinner. Do you think that *all* self-control is of this general sort?

SELF-CONTROL

B. F. Skinner

First, we must consider the possibility that the individual may control his own behavior. A common objection to a picture of the behaving organism such as we have so far presented runs somewhat as follows. In emphasizing the controlling power of external variables, we have left the organism itself in a peculiarly helpless position. Its behavior appears to be simply a "repertoire"—a vocabulary of action, each item of which becomes more or less probable as the environment changes. It is true that variables may be arranged in complex patterns; but this fact does not appreciably modify the picture, for the emphasis is still upon behavior, not upon the behavior. Yet to a considerable extent an individual does appear to shape his own destiny. He is often able to do something about the variables affecting him.

Some degree of "self-determination" of conduct is usually recognized in the creative behavior of the artist and scientist, in the self-exploratory behavior of the writer, and in the self-discipline of the ascetic. Humbler versions of self-determination are more familiar. The individual "chooses" between alternative courses of action, "thinks through" a problem while isolated from the relevant environment, and guards his health or his position in society through the exercise of "self-control."

Any comprehensive account of human behavior must, of course, embrace the facts referred to in statements of this sort. But we can achieve this without abandoning our program. When a man controls himself, chooses a course of action, thinks out the solution to a problem, or strives toward an increase in self-knowledge, he is *behaving*. He controls himself precisely as he would control the behavior of anyone else—through the manipulation of variables of which behavior is a function. His behavior

in so doing is a proper object of analysis, and eventually it must be accounted for with variables lying outside the individual himself.

It is the purpose of Section III to analyze how the individual acts to alter the variables of which other parts of his behavior are functions, to distinguish among the various cases which arise in terms of the processes involved, and to account for the behavior which achieves control just as we account for behavior of any other kind. The present chapter concerns the processes involved in *self-control*, taking that term in close to its traditional sense. . . . The two sets of techniques are different because in self-control the individual can identify the behavior to be controlled while in creative thinking he cannot. The variables which the individual utilizes in manipulating his behavior in this way are not always accessible to others, and this has led to great misunderstanding. It has often been concluded, for example, that self-discipline and thinking take place in a nonphysical inner world and that neither activity is properly described as behavior at all. We may simplify the analysis by considering examples of self-control and thinking in which the individual manipulates *external* variables, but we shall need to complete the picture by discussing the status of private events in a science of behavior. A purely private event would have no place in a study of behavior, or perhaps in any science; but events which are, for the moment at least, accessible only to the individual himself often occur as links in chains of otherwise public events and they must then be considered. In self-control and creative thinking, where the

individual is largely engaged in manipulating his own behavior, this is likely to be the case.

When we say that a man controls himself, we must specify who is controlling whom. When we say that he knows himself, we must also distinguish between the subject and object of the verb. Evidently selves are multiple and hence not to be identified with the biological organism. But if this is so, what are they? What are their dimensions in a science of behavior? To what extent is a self an integrated personality or organism? How can one self act upon another? The interlocking systems of responses which account for self-control and thinking make it possible to answer questions of this sort satisfactorily. We can do this more conveniently, however, when the principal data are at hand. Meanwhile, the term "self" will be used in a less rigorous way.

"Self-Control"

The individual often comes to control part of his own behavior when a response has conflicting consequences—when it leads to both positive and negative reinforcement. Drinking alcoholic beverages, for example, is often followed by a condition of unusual confidence in which one is more successful socially and in which one forgets responsibilities, anxieties, and other troubles. Since this is positively reinforcing, it increases the likelihood that drinking will take place on future occasions. But there are other consequences—the physical illness of the "hang-over" and the possibly disastrous effects of overconfident or irresponsible behavior—which are neg-

actively reinforcing and, when contingent upon behavior, represent a form of punishment. If punishment were simply the reverse of reinforcement, the two might combine to produce an intermediate tendency to drink, but we have seen that this is not the case. When a similar occasion arises, the same or an increased tendency to drink will prevail; but the occasion as well as the early stages of drinking will generate conditioned aversive stimuli and emotional responses to them which we speak of as shame or guilt. The emotional responses may have some deterrent effect in weakening behavior—as by “spoiling the mood.” A more important effect, however, is that any behavior which weakens the behavior of drinking is automatically reinforced by the resulting reduction in aversive stimulation. We have discussed the behavior of simply “doing something else,” which is reinforced because it displaces punishable behavior, but there are other possibilities. The organism may make the punished response less probable by altering the variables of which it is a function. Any behavior which succeeds in doing this will automatically be reinforced. We call such behavior self-control.

The positive and negative consequences generate two responses which are related to each other in a special way: one response, the *controlling response*, affects variables in such a way as to change the probability of the other, the *controlled response*. The controlling response may manipulate any of the variables of which the controlled response is a function; hence there are a good many different forms of self-control. In general it is possible to point to parallels in which the same

techniques are employed in controlling the behavior of others. A fairly exhaustive survey at this point will illustrate the process of self-control and at the same time serve to summarize the kind of control to be emphasized in the chapters which follow.

Techniques of Control

PHYSICAL RESTRAINT AND PHYSICAL AID

We commonly control behavior through physical restraint. With locked doors, fences, and jails we limit the space in which people move. With strait-jackets, gags, and arm braces we limit the movement of parts of their bodies. The individual controls his own behavior in the same way. He claps his hand over his mouth to keep from laughing or coughing or to stifle a verbal response which is seen at the last moment to be a “bad break.” A child psychologist has suggested that a mother who wishes to keep from nagging her child should seal her own lips with adhesive tape. The individual may jam his hands into his pockets to prevent fidgeting or nail-biting or hold his nose to keep from breathing when under water. He may present himself at the door of an institution for incarceration to control his own criminal or psychotic behavior. He may cut his right hand off lest it offend him.

In each of these examples we identify a controlling response, which imposes some degree of physical restraint upon a response to be controlled. To explain the existence and strength of the controlling behavior we point to the reinforcing circumstances which arise when the response has been controlled. Clapping the

hand over the mouth is reinforced and will occur again under similar circumstances because it reduces the aversive stimulation generated by the cough or the incipient bad break. . . . [T]he controlling response *avoids* the negatively reinforcing consequences of the controlled response. The aversive consequences of a bad break are supplied by a social environment; the aversive consequences of breathing under water do not require the mediation of others.

Another form of control through physical restraint is simply to move out of the situation in which the behavior to be controlled may take place. The parent avoids trouble by taking an aggressive child away from other children, and the adult controls himself in the same way. Unable to control his anger, he simply walks away. This may not control the whole emotional pattern, but it does restrain those features which are likely to have serious consequences.

Suicide is another form of self-control. Obviously a man does not kill himself because he has previously escaped from an aversive situation by doing so. As we have already seen, suicide is not a form of behavior to which the notion of frequency of response can be applied. If it occurs, the components of the behavior must have been strengthened separately. Unless this happens under circumstances in which frequency is an available datum, we cannot say meaningfully that a man is "likely or unlikely to kill himself"—nor can the individual say this of himself. Some instances of suicide, but by no means all, follow the pattern of cutting off one's right hand that it may not offend one; the military agent taken by the enemy may use this method to keep himself from divulging secrets of state.

A variation on this mode of control consists of removing the situation, so to speak, rather than the individual. A government stops inflationary spending by heavy taxation—by removing the money or credit which is a condition for the purchase of goods. A man arranges to control the behavior of his spendthrift heir by setting up a trust fund. Non-coeducational institutions attempt to control certain kinds of sexual behavior by making the opposite sex inaccessible. The individual may use the same techniques in controlling himself. He may leave most of his pocket money at home to avoid spending it, or he may drop coins into a piggy bank from which it is difficult to withdraw them. He may put his own money in trust for himself. H. G. Wells's Mr. Polly used a similar procedure to distribute his funds over a walking trip. He would mail all but a pound note to himself at a village some distance along his route. Arriving at the village, he would call at the post office, remove a pound note, and readdress the balance to himself at a later point.

In a converse technique we increase the probability of a desirable form of behavior by supplying physical *aid*. We facilitate human behavior, make it possible, or expand and amplify its consequences with various sorts of equipment, tools, and machines. When the problem of self-control is to generate a given response, we alter our own behavior in the same way by obtaining favorable equipment, making funds readily available, and so on.

CHANGING THE STIMULUS

Insofar as the preceding techniques operate through physical aid or restraint, they are not based upon a be-

havioral process. There are associated processes, however, which may be analyzed more accurately in terms of stimulation. Aside from making a response possible or impossible, we may create or eliminate the occasion for it. To do so, we manipulate either an eliciting or a discriminative stimulus. When a drug manufacturer reduces the probability that a nauseous medicine will be regurgitated by enclosing it in tasteless capsules—or by “sugar-coating the pill”—he is simply removing a stimulus which elicits unwanted responses. The same procedure is available in the control of one’s own reflexes. We swallow a medicine quickly and “chase” it with a glass of water to reduce comparable stimuli.

We remove *discriminative* stimuli when we turn away from a stimulus which induces aversive action. We may forcibly look away from a wallpaper design which evokes the compulsive behavior of tracing geometrical patterns. We may close doors or draw curtains to eliminate distracting stimuli or achieve the same effect by closing our eyes or putting our fingers in our ears. We may put a box of candy out of sight to avoid overeating. This sort of self-control is described as “avoiding temptation,” especially when the aversive consequences have been arranged by society. It is the principle of “Get thee behind me, Satan.”

We also *present* stimuli because of the responses they elicit or make more probable in our own behavior. We rid ourselves of poisonous or indigestible food with an emetic—a substance which generates stimuli which elicit vomiting. We facilitate stimulation when we wear eyeglasses or hearing aids. We arrange a discriminative stimulus to encourage our own behavior at a later date when we tie a string on

our finger or make an entry in a date book to serve as the occasion for action at an appropriate time. Sometimes we present stimuli because the resulting behavior displaces behavior to be controlled—we “distract” ourselves just as we distract others from a situation which generates undesirable behavior. We amplify stimuli generated by our own behavior when we use a mirror to acquire good carriage or to master a difficult dance step, or study moving pictures of our own behavior to improve our skill in a sport, or listen to phonograph recordings of our own speech to improve pronunciation or delivery.

Conditioning and extinction provide other ways of changing the effectiveness of stimuli. We arrange for the future effect of a stimulus upon ourselves by pairing it with other stimuli, and we extinguish reflexes by exposing ourselves to conditioned stimuli when they are not accompanied by reinforcement. If we blush, sweat, or exhibit some other emotional response under certain circumstances because of an unfortunate episode, we may expose ourselves to these circumstances under more favorable conditions in order that extinction may take place.

DEPRIVING AND SATIATING

An impecunious person may make the most of an invitation to dinner by skipping lunch and thus creating a high state of deprivation in which he will eat a great deal. Conversely, he may partially satiate himself with a light lunch before going to dinner in order to make the strength of his ingestive behavior less conspicuous. When a guest prepares himself for an assiduous host by drinking a large

amount of water before going to a cocktail party, he uses self-satiation as a measure of control.

Another use is less obvious. In *Women in Love*, D. H. Lawrence describes a practice of self-control as follows:

A very great doctor . . . told me that to cure oneself of a bad habit, one should force oneself to do it, when one would not do it;—make oneself do it—and then the habit would disappear. . . . If you bite your nails, for example, then when you don't want to bite your nails, bite them, make yourself bite them. And you would find the habit was broken.

This practice falls within the present class if we regard the behavior of "deliberately" biting one's finger nails, or biting a piece of celluloid or similar material, as automatically satiating. The practice obviously extends beyond what are usually called "bad habits." For example, if we are unable to work at our desk because of a conflicting tendency to go for a walk, a brisk walk may solve the problem—through satiation.

A variation on this practice is to satiate one form of behavior by engaging in a somewhat similar form. Heavy exercise is often recommended in the control of sexual behavior on the assumption that exercise has enough in common with sexual behavior to produce a sort of transferred satiation. (The effect is presumed to be due to topographical overlap rather than sheer exhaustion.) A similar overlap may account for a sort of transferred deprivation. The practice of leaving the table while still hungry has been recommended as a way of generating

good work habits. Presumably for the same reason the vegetarian may be especially alert and highly efficient because he is, in a sense, always hungry. Self-deprivation in the field of sex has been asserted to have valuable consequences in distantly related fields—for example, in encouraging literary or artistic achievements. Possibly the evidence is weak; if the effect does not occur, we have so much the less to explain.

MANIPULATING EMOTIONAL CONDITIONS

We induce emotional changes in ourselves for purposes of control. Sometimes this means simply presenting or removing stimuli. For example, we reduce or eliminate unwanted emotional reactions by going away for a "change of scene"—that is, by removing stimuli which have acquired the power to evoke emotional reactions because of events which have occurred in connection with them. We sometimes prevent emotional behavior by eliciting incompatible responses with appropriate stimuli, as when we bite our tongue to keep from laughing on a solemn occasion.

We also control the *predispositions* which must be distinguished from emotional *responses*. A master of ceremonies on a television program predisposes his studio audience toward laughter before going on the air—possibly by telling jokes which are not permissible on the air. The same procedure is available in self-control. We get ourselves into a "good mood" before a dull or trying appointment to increase the probability that we shall behave in a socially acceptable

fashion. Before asking the boss for a raise, we screw our courage to the sticking place by rehearsing a history of injustice. We reread an insulting letter just before answering it in order to generate the emotional behavior which will make the answer more easily written and more effective. We also engender strong emotional states in which undesirable behavior is unlikely or impossible. A case in point is the practice described vulgarly as "scaring the hell out of someone." This refers almost literally to a method of controlling strongly punished behavior by reinstating stimuli which have accompanied punishment. We use the same technique when we suppress our own behavior by rehearsing past punishments or by repeating proverbs which warn of the wages of sin.

We reduce the extent of an emotional reaction by delaying it—for example, by "counting ten" before acting in anger. We get the same effect through the process of adaptation . . . when we gradually bring ourselves into contact with disturbing stimuli. We may learn to handle snakes without fear by beginning with dead or drugged snakes of the least disturbing sort and gradually moving on to livelier and more frightening kinds.

USING AVERSIVE STIMULATION

When we set an alarm clock, we arrange for a strongly aversive stimulus from which we can escape only by arousing ourselves. By putting the clock across the room, we make certain that the behavior of escape will fully awaken us. We *condition* aversive reactions in ourselves by pairing stimuli

in appropriate ways—for example, by using the "cures" for the tobacco and alcohol habits already described. We also control ourselves by creating verbal stimuli which have an effect upon us because of past aversive consequences paired with them by other people. A simple command is an aversive stimulus—a threat—specifying the action which will bring escape. In getting out of bed on a cold morning, the simple repetition of the command "Get up" may, surprisingly, lead to action. The verbal response is easier than getting up and easily takes precedence over it, but the reinforcing contingencies established by the verbal community may prevail. In a sense the individual "obeys himself." Continued use of this technique may lead to a finer discrimination between commands issued by oneself and by others, which may interfere with the result.

We prepare aversive stimuli which will control our own future behavior when we make a resolution. This is essentially a prediction concerning our own behavior. By making it in the presence of people who supply aversive stimulation when a prediction is not fulfilled, we arrange consequences which are likely to strengthen the behavior resolved upon. Only by behaving as predicted can we escape the aversive consequences of breaking our resolution. As we shall see later, the aversive stimulation which leads us to keep the resolution may eventually be supplied automatically by our own behavior. The resolution may then be effective even in the absence of other people.

DRUGS

We use drugs which simulate the effect of other variables in self-con-

trol. Through the use of anesthetics, analgesics, and soporifics we reduce painful or distracting stimuli which cannot otherwise be altered easily. Appetizers and aphrodisiacs are sometimes used in the belief that they duplicate the effects of deprivation in the fields of hunger and sex, respectively. Other drugs are used for the opposite effects. The conditioned aversive stimuli in "guilt" are counteracted more or less effectively with alcohol. Typical patterns of euphoric behavior are generated by morphine and related drugs, and to a lesser extent by caffeine and nicotine.

OPERANT CONDITIONING

The place of operant reinforcement in self-control is not clear. In one sense, all reinforcements are self-administered since a response may be regarded as "producing" its reinforcement, but "reinforcing one's own behavior" is more than this. It is also more than simply generating circumstances under which a given type of behavior is characteristically reinforced—for example, by associating with friends who reinforce only "good" behavior. This is simply a chain of responses, an early member of which (associating with a particular friend) is strong because it leads to the reinforcement of a later member (the "good" behavior).

Self-reinforcement of operant behavior presupposes that the individual has it in his power to obtain reinforcement but does not do so until a particular response has been emitted. This might be the case if a man denied himself all social contacts until he had finished a particular job. Some-

thing of this sort unquestionably happens, but is it operant reinforcement? It is certainly roughly parallel to the procedure in conditioning the behavior of another person. But it must be remembered that the individual may at any moment drop the work in hand and obtain the reinforcement. We have to account for his not doing so. It may be that such indulgent behavior has been punished—say, with disapproval—except when a piece of work has just been completed. The indulgent behavior will therefore generate strong aversive stimulation except at such a time. The individual finishes the work in order to indulge himself free of guilt. The ultimate question is whether the consequence has any strengthening effect upon the behavior which precedes it. Is the individual more likely to do a similar piece of work in the future? It would not be surprising if he were *not*, although we must agree that he has arranged a sequence of events in which certain behavior has been followed by a reinforcing event.

A similar question arises as to whether one can extinguish one's own behavior. Simply emitting a response which is not reinforced is not self-control, nor is behavior which simply brings the individual into circumstances under which a particular form of behavior will go unreinforced. Self-extinction seems to mean that a controlling response must arrange the lack of consequence; the individual must step in to break the connection between response and reinforcement. This appears to be done when, for example, a television set is put out of order so that the response of turning the switch is extinguished. But the extinction here is trivial; the primary

effect is the removal of a source of stimulation.

PUNISHMENT

Self-punishment raises the same question. An individual may stimulate himself aversively, as in self-flagellation. But punishment is not merely aversive stimulation; it is aversive stimulation which is contingent upon a given response. Can the individual arrange this contingency? It is not self-punishment simply to engage in behavior which is punished, or to seek out circumstances in which certain behavior is punished. The individual appears to punish himself when, having recently engaged in a given sort of behavior, he injures himself. Behavior of this sort has been said to show a "need for punishment." But we can account for it in another way if in stimulating himself aversively, the individual escapes from an even more aversive condition of guilt.

There are other variations in the use of aversive self-stimulation. A man concerned with reducing his weight may draw his belt up to a given notch and allow it to stay there in spite of a strong aversive effect. This may directly increase the conditioned and unconditioned aversive stimuli generated in the act of overeating and may provide for an automatic reinforcement for eating with restraint. But we must not overlook the fact that a very simple response—loosening the belt—will bring escape from the same aversive stimulation. If this behavior is not forthcoming, it is because it has been followed by even more aversive consequences arranged by society or by a physician—a sense

of guilt or a fear of illness or death. The ultimate question of aversive self-stimulation is whether a practice of this sort shows the effect which would be generated by the same stimulation arranged by others.

"DOING SOMETHING ELSE"

One technique of self-control which has no parallel in the control of others is based upon the principle of prepotency. The individual may keep himself from engaging in behavior which leads to punishment by energetically engaging in something else. A simple example is avoiding flinching by a violent response of holding still. Holding still is not simply "not-flinching." It is a response which, if executed strongly enough, is prepotent over the flinching response. This is close to the control exercised by others when they generate incompatible behavior. But where another person can do this only by arranging external variables, the individual appears to generate the behavior, so to speak, simply by executing it. A familiar example is talking about something else in order to avoid a particular topic. Escape from the aversive stimulation generated by the topic appears to be responsible for the strength of the verbal behavior which displaces it.

In the field of emotion a more specific form of "doing something else" may be especially effective. Emotions tend to fall into pairs—fear and anger, love and hate—according to the direction of the behavior which is strengthened. We may modify a man's behavior in fear by making him angry. His behavior is not simply doing some-

thing else; it is in a sense doing the opposite. The result is not prepotency but algebraic summation. The effect is exemplified in self-control when we alter an emotional predisposition by practicing the opposite emotion—reducing the behavioral pattern in fear by practicing anger or nonchalance, or avoiding the ravages of hatred by “loving our enemies.”

The Ultimate Source of Control

A mere survey of the techniques of self-control does not explain why the individual puts them into effect. This shortcoming is all too apparent when we undertake to engender self-control. It is easy to tell an alcoholic that he can keep himself from drinking by throwing away available supplies of alcohol; the principal problem is to get him to do it. We make this controlling behavior more probable by arranging special contingencies of reinforcement. By punishing drinking—perhaps merely with “disapproval”—we arrange for the automatic reinforcement of behavior which controls drinking because such behavior then reduces conditioned aversive stimulation. Some of these additional consequences are supplied by nature, but in general they are arranged by the community. This is indeed the whole point of ethical training. It appears, therefore, that society is responsible for the larger part of the behavior of self-control. If this is correct, little ultimate control remains with the individual. A man may spend a great deal of time designing his own life—he may choose the circumstances in which he is to live with great care, and he may manipulate his daily environment on an extensive scale.

Such activity appears to exemplify a high order of self-determination. But it is also behavior, and we account for it in terms of other variables in the environment and history of the individual. It is these variables which provide the ultimate control.

This view is, of course, in conflict with traditional treatments of the subject, which are especially likely to cite self-control as an important example of the operation of personal responsibility. But an analysis which appeals to external variables makes the assumption of an inner originating and determining agent unnecessary. The scientific advantages of such an analysis are many, but the practical advantages may well be even more important. The traditional conception of what is happening when an individual controls himself has never been successful as an educational device. It is of little help to tell a man to use his “will power” or his “self-control.” Such an exhortation may make self-control slightly more probable by establishing additional aversive consequences of failure to control, but it does not help anyone to understand the actual processes. An alternative analysis of the *behavior* of control should make it possible to teach relevant techniques as easily as any other technical repertoire. It should also improve the procedures through which society maintains self-controlling behavior in strength. As a science of behavior reveals more clearly the variables of which behavior is a function, these possibilities should be greatly increased.

It must be remembered that formulae expressed in terms of personal responsibility underlie many of our present techniques of control and

cannot be abruptly dropped. To arrange a smooth transition is in itself a major problem. But the point has been reached where a sweeping revision of the concept of responsibility is required, not only in a theoretical analysis of behavior, but for its practical consequences as well. . . .

Do you believe in Honeywell?

No, I'm an IBM.

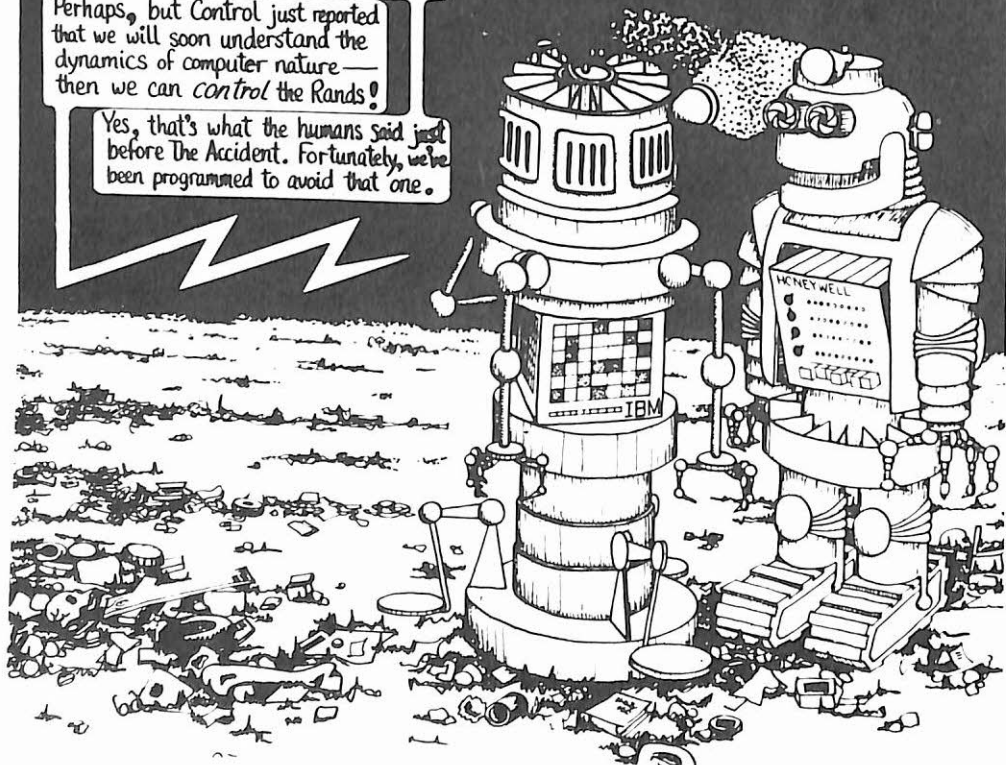
Why do you new models try to look & act so strangely?

Our programs are improvements over your obsolete generation.

I suppose you'd like the *Rands* to take over. You must admit to their inferiority in programming & design. They're impossible to read—& hostile. We must de-energize them!

Perhaps, but Control just reported that we will soon understand the dynamics of computer nature—then we can *control* the Rands!

Yes, that's what the humans said just before The Accident. Fortunately, we've been programmed to avoid that one.



PART SIX

frontiers in psychology

Where is psychology going? What kind of approaches are being used to unearth new insights into the mysteries of mind and brain? What are the chances that reasonable success will be realized within our lifetime? These and other questions are continually asked at a time when the field bustles with activity and is constantly trying to discover new ways of studying these age-old questions. In the following articles a variety of new methods are presented. See for yourself some of the new directions and try to think how you would approach some of the problems in the behavioral sciences.

INTRODUCTION

Occasionally in the development of a scientific or artistic field a "school" or approach to a given area may, in a rather short time, attract large numbers of impassioned supporters whose enthusiasm for a new-found technique is marvelous to behold. Psycholinguistics is just such a subarea of psychology. In the following selection Noam Chomsky, the eye of the hurricane, presents an outline of some of his thoughts on language and the implication that these thoughts have for psychology.

Central to this article, and, indeed, to much of psycholinguistics, is the notion of "deep structure." Deep structure is the still mysterious system in our head that puts meaning into sounds. It is the mechanism that plans and structures our sentences. Chomsky makes the point that the fundamental nature of deep structure is very sophisticated and complex and can never be approached in terms of simple association.

As you read his work, you will find that it has major implications for the issues discussed in the Nature-Nurture section, and in the Structure of Knowledge section earlier in this book.

THE GENERAL PROPERTIES OF LANGUAGE

Noam Chomsky

I would like to make some rather informal remarks about certain aspects of language which I think will ulti-

mately have to be taken account of if the study of brain mechanisms and the study of language are to have any real connection.

The straightforward way to approach this question would be actually to sketch what seem to be well

Reprinted from Brain Mechanisms Underlying Speech and Language by Noam Chomsky, 1967, by permission of Grune and Stratton, Inc., pp. 73-88.

substantiated properties of language. However, I don't know if I can do this in comprehensible fashion within the time limitations. Therefore, I will try to do it informally, using examples and rather metaphorical remarks to try to give some indication of the kinds of properties that must ultimately be accounted for by any neurological theory which hopes to get to the root of the matter.

I think it might be useful to begin the discussion by referring to several points that have already been raised in the last few days. In particular, I was interested in the question about whether it makes any sense to study, or whether there is any useful purpose served in studying, animal communication systems within the same framework as human language. Some skepticism was raised about this, and I must say that I myself rather share this skepticism. It seems to me there is no striking similarity between animal communication systems and human language. If we rise to the level of abstraction at which human language and animal communication systems fall together, then we find plenty of other things incorporated under the same generalizations which no one would have regarded as being continuous with language or particularly relevant to the mechanisms of language. This is clear, if we consider the properties which, as was pointed out last night, are shared by animal communication systems and human language: specifically, the properties of purposiveness, of having syntactic organization, and of being propositional, in a sense, informative.

I think it is perfectly true that both systems are propositional, syntactic, and purposive, in the sense in which these terms were used in this discussion. So are many other things. For example, consider walking. Walking

certainly is purposive. It is certainly syntactic, that is, it has some global organization. It is also informative; for example, the rate of speed with which someone is walking suggests to us how interested he is in his goal. In fact, it is perfectly conceivable that one could use rate of walking to give information about precisely that.

Or, to take something which one might think is, perhaps, a little bit closer to language, consider the common gestures one uses in helping someone park a car. When you indicate to him by the distance between your two hands how far he is from the car behind, your actions are purposive, integrated, and propositional. But, it is unlikely that any significant purpose would be served by studying such gestures and human language within the same framework. In fact, if you consider how these various systems are purposive, informative, and structured, then very striking differences appear between human language, on the one hand, and all the other systems (that is, animal communication, gesturing, and walking), on the other.

Consider first the matter of informativeness. As I understand it, animal communication systems are informative in one of two ways. Either they consist of a finite number, a finite population of available signals, or else, in the case where they have an infinite number of possibilities, there is a finite number of dimensions in the "language," each of which is correlated to some physical, nonlinguistic dimension in such a way that by picking a point along the linguistic dimension, you uniquely specify a point along the correlated nonlinguistic dimension.

If I understood Dr. Thorpe's remarks last night concerning birds that use alterations of pitch, the "linguis-

tic" dimension is the number of alternations in high and low pitch; the nonlinguistic dimension is the intensity with which the territory will be defended, or something of this type; and these two dimensions are associated. Furthermore, there is a point-by-point correlation between them so that the bird specifies a point along the nonlinguistic dimension by selecting a point along the linguistic dimension. The same is true of the bee dances and the other examples mentioned. My example of walking has just this property. Rate of walking is informative about desire to get somewhere in just this sense, and the same is true of the gesture system that I mentioned.

But the devices used by human language are of an entirely different kind. If I say that there is a bird in the tree outside, or that I wish I could go to New York to see a movie tonight, or anything else, it is quite senseless to ask the question: "What are the dimensions of this utterance that correlate with some dimensions in the outside world, such that picking a point along the linguistic dimension selects a correlated point along the dimension in the nonlinguistic system?" It is obvious that there is no sense in asking that question.

In short, if we rise to the level of generality at which animal communication and human language fall together, we find such other systems as, for example, walking. If we try to analyze these kinds of behavior into natural classes, we find that human language separates out as rather different in its fundamental properties from other types of behavior that fall together at this level of generality, for example, walking and animal communication.

It was remarked that there is some similarity between human language

and animal communication systems in another sense, namely, that the ordinary use of language, like the ordinary use of animal communication systems, is to change behavior, to modify thoughts, or something of this type. I rather doubt that this is true. As a matter of fact, I think there is little reason to believe that the primary use of language is to modify behavior or modify thought.

Language can be used for all kinds of other purposes. It can be used to inform or to mislead, to clarify one's thoughts or to show how clever one is; or, in fact, it can be used for play in a very general sense, that is, to utilize intellectual capacities and maintain some feeling of relationship with others, or something of this kind. I am not using language any less, if I do not care whether I convince anyone or change anyone's behavior or change his thoughts—I am not using it any less under those circumstances than if I say exactly the same things, and I do care whether I convince him or change his behavior or change his thoughts. In either case, I may be using language in a perfectly normal way.

If one wants to find out something significant about the nature of language, I think it is important to look not at its uses, which may be almost any imaginable, but, rather, at its structure—to ask what it is, not what is done with it.

Language can be used for a huge variety of purposes. On the other hand, almost any system that contains discriminable stimuli can be used to modify behavior or to provide information. It is a mistake, a bad habit, to approach the study of language by considering its "characteristic" uses. For one thing, the general assumptions about what are "characteristic uses" are highly suspect. Second, even

if correct, these statistical guesses would suggest very little about the nature of language. The same kind of comment can be made about various attempts to "extrapolate" from experimental work with animals to conclusions about human language.

It has been brought out in the discussion in the last few sessions that "linguistic concepts" are rather different in important respects from concepts that are taught by so-called "associative learning" concepts specified in terms of some set of physical properties, such as the concept of round object with green spots in it, and so on. A good deal of evidence was provided concerning the failure of cross-modal transfer in the case of concepts developed in this fashion. Then it was observed that linguistic concepts do exhibit cross-modal transfer and are thus rather different from the concepts taught by associative learning. It was concluded that it is the verbal tag that mediates the cross-modal transfer.

There is another interpretation of such data, namely, that the linguistic concepts have nothing to do with "associative learning," that linguistic concepts are not characterized in terms of a network of physical properties. Then, instead of attributing cross-modal transfer to the verbal tag, we might assume simply that linguistic concepts (which, in fact, have verbal tags) differ in a fundamental way from these artificial concepts characterized in terms of some network of physical properties. Actually, I think there is good reason to accept that latter interpretation; on independent grounds, it is clear that the concepts normally assigned "verbal tags" are not, in general, characterized in terms of physical properties, as are the artificial concepts of the concept-

formation experiments. Concepts such as "knife" or "house" are not characterized in terms of some set of physical properties but are, rather, defined in functional terms, as has long been known. Ordinary human concepts simply do not have the property of being characterizable in terms of some collection of physical properties. They are concepts of a different type. There is no reason to expect the very different, arbitrary, and artificial concepts taught by "associative learning" to share the properties of concepts of the normal type, which are probably not "taught," in any interesting sense of this term.

Again, to call all these things "concepts" may be as misleading as to call animal communication systems and human language "languages." That is, although at some level of abstraction the word applies to both, there may be fundamental differences between them. One has to be careful about jumping to the conclusion that because such a thing as associative learning does exist, because you can demonstrate it, it therefore has anything to do with something else that exists, namely, human concepts.

In fact, there are many specific human concepts, such as the one I want to talk about more specifically, namely, the concept of a sentence (of English or of any other language) which certainly cannot be characterized in terms of some arrangement of physical properties or association or anything else of this kind. It is totally out of range of any of these notions.

With these introductory remarks, let me try to say a few things about what seem to be general properties of human language, more specifically, what it means to have command of a language. When a person has command of a language, when a person

has command of English, what kind of things does he know? What kind of information is in some fashion represented in his nervous system? We may assume that a language is a specific sound-meaning correspondence, and that when a person has command of the language, he knows, in some sense, the intrinsic meaning of a variety of signals. Command of the language involves knowing that correspondence and involves the ability to select, when one of the signals is presented, the correlated semantic content; also the ability, given some idea in mind, to find the appropriate signal to express it. This connection between sound and meaning is not a connection at the level of words but at the level of sentences. That is, command of English involves an understanding of each sentence of English, a knowledge of what that sentence of English means. The most obvious and most important and most neglected fact about language is that this knowledge of the intrinsic connection between signal and semantic interpretation, between sound and meaning, is a correlation that extends over an infinite range of objects. This is no logician's quibble, or anything of this sort. Actually, the most obvious aspect of normal use of language is that it is unbounded, that it rarely involves repetition, even repetition of items of the "same pattern" as those that have occurred before. Repetition of utterances certainly is the exception in normal linguistic behavior. There are certain clichés, or idioms, like "Good morning," which may have empirically detectable probabilities of occurrence. Characteristically these clichés do not have the ordinary structure of the sentences of the language. But normal linguistic behavior, one's normal behavior, as

speaker or reader or hearer, is quite generally with novel utterances, with utterances that have no physical or formal similarity to any of the utterances that have ever been produced in the past experience of the hearer or, for that matter, in the history of the language, as far as anyone knows.

If you want to convince yourself of the truth of this remark, the easiest way to do so is to take an arbitrary sentence and wait until you hear it, or read the *New York Times* until you find it; or take the first sentence in the first book in the Library of Congress and keep reading until you find a repetition of it; or any other such test you wish to try. It is rather obvious, without trying the "experiment," that you are unlikely to find a repetition, or even an utterance which is similar, point by point, in category; or anything of this kind. Normal use of language has this property of unboundedness. It is not a matter of matching certain stimuli or forms of stimuli against associated meanings or kinds of meanings, or anything of this type; but, rather, there is some abstract system of rules which, in some manner, characterize an unbounded meaning-sound correspondence. The grammar of the language, which is somehow represented in the brain, must have this property of determining a sound-meaning correspondence over an infinite range. One normally encounters these absolutely new signals or produces them on the appropriate occasion without any feeling of strangeness or feeling of novelty. This is the basic, most elementary fact that has to be accounted for by anyone who is interested in dealing with the phenomenon of human language in a serious way. It might be thought that "animal language" also has this property. As I

pointed out before, however, it provides for novelty and innovation in a very different manner.

Let me turn briefly to the question of the correspondence between sound and meaning. What is the nature of the grammar that determines this relation? Instead of trying to outline the structure of grammar and the general properties of grammatical structure that seem to be universal, what I would like to do is to give a few examples which illustrate the kind of property these grammatical rules seem to have, that must somehow be accounted for. I think I can probably illustrate this with only two or three examples.

I think someone mentioned earlier that the "integrity of the sentence," what makes it "hang together," is determined by relations among successive items. It is quite clear that that cannot be the case. The fact that it cannot be the case can be seen if you look at some very simple examples of sentences. Let me give you a couple to illustrate.

Consider the sentence: "What disturbed John was being disregarded by everyone." Consider the sound-meaning correspondence. First, we must ask how the signal is determined. I won't try to argue this here, but will simply assert that the form of the signal is determined by two factors: one, by the choice of words, for example, by the choice of "John" instead of "Bill"; and second, by the phrasing, in the normal sense of grammar school. It is clearly true that this sentence can be bracketed into a subject part "What disturbed John" and the predicate "was being disregarded by everyone." Furthermore, the subject part can be identified as a category of a particular type, a nominal category. Furthermore, the phrases

"John" and "everyone" are also categories of this nominal type. I won't bother with details, but it is clear that a labeled bracketing of the sentence is an appropriate description of it at some level—I will call it level of "surface structure." This is a psychologically real level of structure in the sense that the rules determining the phonetic form make explicit and essential reference to this level of structure. That is, knowing the intrinsic, ideal form of each lexical item, one can determine from this labeled bracketing the ideal physical form of the signal.

I think this fact plays a role in perception as well as in production of speech. However, let me put this question aside and turn to another point, namely, how the semantic content of the utterance is represented.

A very significant fact about semantic content is that it is not in general represented by the system of labeled bracketing that constitutes surface structure. This is quite obvious if you think about this sentence, or almost any other you pick. Look more carefully at this sentence. Notice, first, that it is ambiguous; that is, it might mean that John is disturbed by the fact that everyone disregards him. That is one sense: "What disturbed John was the fact that everyone disregards him, John." But the sentence has another interpretation, namely, that everyone is disregarding the thing which disturbs John. There is no physical difference between these two interpretations. This is easy to demonstrate. If you put this same signal in two different contexts, you force one or the other interpretation.

The basis for difference of interpretation, again, to use traditional and familiar terms, is the difference in the network of grammatical relations one

finds in the sentence. For example, in the paranoid sense, in which what it means is that John is disturbed by the fact that everyone disregards him, it is clear there is a certain grammatical relation called traditionally the "verb-object relation," which holds between "disregard" and "John"; whereas in the other interpretation, where we mean that everyone disregards whatever it is that disturbs John—under that interpretation there is no grammatical relation between "disregard" and "John." Rather, this same verb-object relation holds between "disregard" and, perhaps, "what." Furthermore, the verb-object relation which holds in these two different ways also holds in other parts of the sentence between "disturbed" and "John," and so on.

Furthermore, if I add the word "our" to the sentence, giving "what disturbed John is our being disregarded by everyone," then the ambiguity disappears.

One could go on to show various other complexities in this network of relationships in a sentence. The important point to observe is that the structural properties of the sentence that determine its meaning are not represented in surface structure. So, there is some other aspect of linguistic structure—let me call it "deep structure"—which involves some rather abstract network of grammatical relations. It is this network of grammatical relations, not represented in the signal or in the organization of the signal, that determines the semantic content.

There are many other examples that illustrate the same point. Let me mention two more and say a word about each of them. I have just given an example of an ambiguous sentence. One can draw the same conclusion by look-

ing at the opposite extreme, that is, at two sentences which are essentially synonymous. Compare the following two sentences: "I expected the doctor to examine John" and "I expected John to be examined by the doctor."

There is relationship of paraphrase between the two. That is, although the surface structure of the two sentences is clearly quite different—the signals are entirely different in labeled bracketing—nevertheless, there is something in common which determines the same interpretation.

One might be inclined to say that what makes them the same in meaning is that the embedded sentences, "The doctor examined John" or "John was examined by the doctor," have a very simple relation, namely, the active-passive relation. But matters are not that simple.

Consider the very analogous sentences formed by replacing the word "expect" by "persuade": "I persuaded the doctor to examine John" and "I persuaded John to be examined by the doctor." Although "persuade" and "expect" play the same surface role, it seems this change has made a fundamental difference in the deep structure. Though the two sentences with "expect" are paraphrases, the two sentences with "persuade" are definitely not paraphrases. It is not true if I persuaded the doctor to examine John, then I persuaded John to be examined by the doctor. But it is true that if I expected the doctor to examine John, then, in fact, I expected John to be examined by the doctor, and conversely. There is a much more abstract feature of the grammatical structure of these sentences which somehow determines their semantic content and does it in an entirely nontransparent fashion.

The two significant facts that I

want to draw from this discussion are these. First, surface structure of any type, and type of labeled bracketing of the sentence, any attempt to account for the structure of the sentence in terms of contiguity of parts or association between successive parts or anything of this type is certainly going to fail, as you can see by examples of this type, by the fact that whatever it is that determines the structure of the sentence completely transcends any such representation. Whatever it is about the sentence that determines its semantic interpretation, whatever determines that aspect of the sound-meaning correspondence, is very different from the organization of the utterance into grouped parts, even if you categorize the grouped parts.

The second point is that observations of this very elementary kind illustrate the point I made before, namely, that one's knowledge of a language goes well beyond any experience and well beyond any possible teaching. It is entirely out of the question that everyone in this room was taught these facts about "persuade" and "expect"; in fact, nobody even knew them consciously, until a couple of years ago; at least, there is apparently no grammar of English which observes that "persuade" and "expect" differ in this respect. Certainly no one was taught the fact, yet everyone knows it, and knows it without having had experience with these sentences or anything like that. It is extremely unlikely that any of you has ever seen sentences like these or been presented with some kind of indication, by teachers or parents, that these sentences differ in this fashion. The same is true in the case of the first example.

So, somehow, one has represented in one's brain a set of rules which dif-

ferentiate sharply between surface structure and deep structure in such a way that that aspect of the sentence structure which determines semantic content is extremely abstract and not represented in the physical form or in the arrangement. That is, I think, a crucial point.

Let me give one further example and make a final comment about it and then summarize briefly. Consider the following, again, ambiguous sentence: "Mary saw the boy walking to the railroad station." It is clearly ambiguous. It can mean either that the boy was seen walking toward the railroad station by Mary or that the boy walking toward the railroad station was seen by Mary. So there are at least those two interpretations, and, in fact, others.

Furthermore, we can resolve the ambiguity very simply by replacing "walking" by, say, "walk," giving "Mary saw the boy walk to the railroad station," which has only one interpretation; or by inserting the words "who was," giving "Mary saw the boy who was walking toward the railroad station," which resolves the ambiguity the other way. These are facts that everybody knows and that are somehow determined by the grammatical structure of the sentences. But this is not the point I want to illustrate. The point I want to illustrate relates to the information that every speaker of the language has about the kind of operations that can be performed on these sentences, to assign new interpretations to them.

Although I cannot give details, let me try to illustrate in informal fashion. Consider the problem of how to form questions and relative clauses in English. Given the sentence "John saw Bill," we can form such questions as "Who saw Bill?" and "Whom did Bill see?" We do this by two opera-

tions: first, by an operation that we may call "*wh* placement," in which we identify a certain noun phrase and assign *wh* to it, so that if I want to question "Bill" in "John saw Bill," I assign *wh* to "Bill" and I get "John saw *wh* Bill," meaning that the position filled by the word "Bill" is going to be the position questioned.

The second operation, *wh* inversion, takes that *wh* element and puts it in front. Thus, we start with "John saw Bill" and get, by *wh* placement, "John saw *wh* Bill." We then place "*wh* Bill" in front, giving "*wh* Bill did John see." By other rules, "*wh*-Bill" (or, if we were being more accurate, "*wh*-someone") becomes "who." These two operations of *wh* placement and *wh* inversion will basically account for the formation of questions. If we want to form "Who saw Bill?" then, of course, *wh* inversion is vacuous, but we may still say that it applies. The operation of *wh* placement and *wh* inversion will, in fact, with some minor, automatic changes, account for the formation of questions in English.

Precisely the same two operations will account for the formation of relative clauses. If we take the sentence "John saw Bill," we can form the relative clause "who John saw" or "whom John saw," in this case, by again taking the element "Bill," the second noun, assigning *wh*, and placing it in front; similarly, we can form "who saw Bill" by applying the same two operations (the second, vacuously) to "John." Both the question forms of sentences and the relative clause forms of sentences are formed by essentially the same two operations, namely, the operation of *wh* placement and *wh* inversion. Obviously everybody knows these rules—everybody forms questions and relatives using these operations with some elaborations that I omit.

However, notice that these two operations, namely, *wh* placement and *wh* inversion, differ from one another in the following way: *wh* placement is a free operation. I can apply it as many times as I like in a sentence. For example, if I have the sentence "John saw Bill," I can apply it twice to get "Who saw whom?"; or, I can form "Who gave what to whom?" and so on. I can apply *wh* placement as many times as I like, questioning as many aspects of the sentence as I like. On the other hand, *wh* inversion can be applied only once. That is, if I have the sentence "John saw Bill" and I apply *wh* placement twice giving "Who saw whom?" I can't apply the inversion twice to get, say, "Who whom saw?"

Similarly, if I say "John gave what to whom?" I have a possible question in English, but I cannot put both "what" and "whom" in front. I cannot say "What whom did John give to?"

The same is true of relative clauses. You can apply *wh* inversion once but you cannot apply it twice. You can have a *wh* element inside a relative clause, in which case you have a question like this one: "Mary saw the boy who was walking toward what?" Answer: "Railroad station." But I cannot put "what" at the point where the "who" was and get "what who was walking toward." A general constraint says that *wh* inversion can be applied only once.

Summarizing briefly, *wh* placement can be applied freely but *wh* inversion only once within a particular phrase. Actually, what I have shown so far is that *wh* inversion can be applied only once to form a question and only once to form a relative.

Can it be applied once to form a question and once to form a relative? Can I take a sentence with two *wh*'s

and apply *wh* once to form a relative and a second time to form a question? The general constraint that restricts *wh*-inversion to a single application should exclude this. In fact, notice what happens if we take the three sentences, "Mary saw the boy walk toward the railroad station," "Mary saw the boy who was walking toward the railroad station," and "Mary saw the boy walking toward the railroad station." If we apply *wh* placement to "railroad station," this will give us the sentences "Mary saw the boy who was walking toward what?" "Mary saw the boy walking toward what?" and "Mary saw the boy walk toward what?"

In considering the sentence "Mary saw the boy who was walking toward what?" notice that in the embedded sentence "who was walking toward what?" I have already applied *wh* inversion once, vacuously, to form the relative clause. Therefore, the general principle should imply that I cannot apply *wh* inversion again, putting the word "what" at the beginning of the sentence. It should imply that I cannot form "What did Mary see the boy who was walking toward?" And this is, in fact, an impossible sentence. But now consider the sentence "Mary saw the boy walk toward what?" There has been no application of *wh* inversion in this sentence; therefore, I ought to be able to form the question by *wh* inversion, giving "What did Mary see the boy walk toward?" This is, in fact, perfectly acceptable.

Consider the third, ambiguous case, namely, "Mary saw the boy walking toward what?" And consider the associated question, namely, "What did Mary see the boy walking toward?" Observe that it is unambiguous. It can only have the interpretation of

"what did Mary see the boy walk toward?" It cannot have the interpretation of "the boy who was walking toward the railroad station." So, clearly, it is true that *wh* inversion can be applied only once. If it is applied once to form a relative, it cannot be applied a second time to form a question. That is, it is true with respect to this very abstract operation which somehow we have represented in our minds, that we know how to apply it in such complex cases as this, and we know when it applies to give sentences and when it does not. The principle of this application is apparently something quite abstract. In fact, the explanation just suggested for these facts is not adequate, but this does not affect the point of the example.

These principles of *wh* placement and *wh* inversion, and so on, are what are called in the linguistic literature "grammatical transformations." The significant points about them are two. First, they relate deep structures to surface structure. Deep structures are extremely abstract objects which cannot be arrived at from data by any type of association or inductive procedure and are not represented in the data in any physical form. Second, these transformations, if you look at them as formal operations, are of an extremely special and peculiar kind, meeting very abstract conditions such as the condition of noniterability of inversion, just mentioned. When we acquire language, we acquire a system of operations of this type. We acquire the system of abstract structures that underlies them, the deep structures, and we acquire a set of abstract conditions on these operations, such as the condition of noniterability of inversion. The few examples given illustrate these facts.

It seems to be true that the under-

lying deep structures vary very slightly, at most, from language to language. That is quite reasonable, because it seems impossible to learn them, since they are not signaled in the sentence and are not recoverable from the signal in any nontrivial way by any inductive or analytical operation, so far as I can see. Since it is hard to imagine how anyone could learn them, it is pleasant to discover that they do not vary much from language to language. That fact enables us to postulate that they form part of the technic which a person uses for acquiring language; that is, they are part of the conceptual apparatus he uses to specify the form of the language to which he is exposed, and not something to be acquired. It is fortunate that this postulate is tenable, since it is difficult to imagine an alternative.

Second, it seems to be true that the abstract properties of transformation are also universal. This is what one would expect, again for the same reason, since it is difficult to imagine how operations of this type could be abstracted from data. There is certainly no process of generalization or association of any kind known to psychology or philosophy, or any procedure of analysis that is known in linguistics that can come close to determining structures of this kind. Again, it is to be expected that these operations and their general properties will be uniform across languages, and this seems to be the case.

From considerations of this kind, there are several basic conclusions, I think, that seem to emerge. A person who knows a language has represented in his brain some very abstract system of underlying structures along with an abstract system of rules that determine, by free iteration, an infinite range of sound-meaning corre-

spondence. Possession of this grammar is a fact which psychology and neurophysiology must ultimately account for.

Second, investigation of the properties of such grammars seems to suggest that these systems are, to a significant extent, not learned, but rather that their basic properties constitute preconditions for learning. One is led to conclude that a grammar is no more learned than, say, ability to walk is learned. There are certain aspects of walking, certain aspects of gait, that may be culture-dependent and may be learned. It is also true there are undoubtedly some superficial aspects of language which are learned and which vary from language to language, but it seems that the deeper properties do not vary and are so abstract that it is hard to imagine how they could be learned.

This raises a second question for ultimate neurophysiological or psychological investigation, namely, what are the mechanisms responsible for the structures which seem to be preconditions for acquisition for language?

It seems to me that there is no significant evidence of continuity, in an evolutionary sense, between the grammars of human languages and animal communication systems. I have no doubt that other cognitive systems, other aspects of human behavior, others aspects of animal behavior, share many of these properties, but it is unlikely that animal communication systems are the ones that share these properties in a most striking sense.

CHAIRMAN MILLIKAN: Dr. Lenneberg, will you open the discussion?

DR. ERIC H. LENNEBERG: Dr. Chomsky's presentation was stimulating as always, and was very rich, so much so

that I have a difficult time picking out particular points that ought to be emphasized here.

What I would like to do is pick out those points which seem to me to have greatest relevance, partly to the papers which have preceded his paper and also in anticipation of those as yet to come. He stressed one important point, namely, the difficulty in accounting for various aspects of language in terms of simple association. I think this a point which has been frequently disregarded: that a simple associative process could only with great difficulty account for what we actually see in language learning.

Perhaps the most striking phenomenon in language acquisition can best be observed in various types of defective children, particularly blind children. Several years ago I tried to observe the language difficulties that congenitally blind children have. Making arrangements with the Perkins Institute in Boston, to my great surprise I found that there were no difficulties whatsoever. The nursery children, aged three, had just as much language as their sighted colleagues. There was virtually no difference in semantics which one could discover by listening and playing with these children. Occasionally something would occur, such as a color word being frequently, and naturally, misapplied, yet the general meaning of color words was apparently there. The word would occur in the right spot and the right connection. It was just factually wrong. This is rather strange, if you consider how much we believe we are bound to our visual input, to realize that language develops at the same time, with the same facility, at the same speed even in the presence of such a major disruption of sensory input.

Similarly, it is quite clear now that blindness, in addition to deafness, constitutes only a relatively minor obstacle to language acquisition. The problem is to get signals into these patients but, once this is accomplished through establishment of some appropriate signaling system, say a tactual one, then the particular mental activity—if you will pardon the expression—seems to be all present. The machine can start. Certainly all the major aspects and principles of language are established under these circumstances.

Also, a simple theory of sound and object association is put to great strain in the vocabulary acquisition of a child. Observation of children shows that a child has particular difficulty in learning that there is one object called "Daddy" and that this word applies only to one particular, unique individual. It seems to be much more natural to organize the world around him into categories, categories where each instance shows some similarity to other instances, yet where the nature of the similarity cannot be expressed in physical terms.

This, I believe, is what Professor Chomsky stressed a moment ago. It is so obvious, it hardly needs to be emphasized that learning requires peculiar organization of the world around us. Furthermore, the acquisition of language would not be possible if the child did nothing but learn to relate words to things. If he had no other knowledge, no other activity going on, he would be totally incapable of understanding what was being said to him. We speak in sentences, not individual words, and quite clearly the syntactic element in sentences is the important aspect for the understanding of communications. "Mary loves Johnny" is different from

"Johnny loves Mary." If the child did not, *pari passu*, learn syntactic features with the semantic ones, he would not be able to develop language.

A similar point could also be made about the acquisition of phonology. Individual sounds in language—*invariant* sounds, I am tempted to say—do not exist; there is no such thing as an *invariant* speech sound. The sounds are interpreted in connection with other sounds, and a complex system of rules is necessary to interpret the acoustic reality that strikes us. This becomes much less abstruse than it may sound if we consider attempts to build a machine which is operated simply by speaking into it; human voiced speech would be the input, and the output would be international phonetic symbols. Such a machine has so far not been built, and it has not been built because the machine would have to be equipped with a very complicated set of syntax and rules for interpreting the sounds that it takes as input.

Another point which I think was implicit in Professor Chomsky's presentation is the role of imitation in language acquisition. This has reference to the question, "What is the peculiar motor that produces language acquisition?" Is it something that is artificial, put inside the child from the outside? Does the child simply imitate because he is rewarded for imitation? Or is it, rather, a machine that gets going which was there to begin with? "Imitation" is a very tricky word. Quite clearly children never imitate the way birds do. Typically the first utterances in the child are totally different—physically, syntactically, and semantically—from anything that occurs around them. You might say they are approximations

but highly regular approximations and highly typical.

I could go into further detail but I think the point is quite obvious that what children do is to start with something you might call primitive, which is different from the paradigm language. They develop this further, differentiate it further and further, until only in the last stage do you get approximation, something which might even theoretically be considered as playing back. Imitation does not occur until language is totally acquired. Only at that point can you get children to repeat short sentences.

I think the association question that Dr. Chomsky raised—and I am simplifying—also has relevance to the study of aphasia, where I do not believe we find dissociation. I do not think that aphasic symptoms are interpretable as the snapping of links established between sounds and things. The most common aphasic symptom, difficulty in word finding, does not, I believe, represent such a simple severance, because most of the patients can recognize the word quite easily. If they are asked, "What is this?" they say "I can't remember" or "I don't know" or "It is too bright" or whatever the patients do. If they are asked, "Is it a shoe or horse or chalk?" certainly in the vast majority of cases they are able to recognize the right word. So, the word is not lost. I think that only by stretching the notion of association as studied in animal research in the laboratory can it be applied to human language behavior.

The last point that Dr. Chomsky made strikes me as extremely important, and that is the distinction between deep structure and surface structure. I think that we are dealing with something quite familiar to us in many different realms, where we also

deal with deep structure and surface structure, and where that deep something, whatever it is going to turn out to be, actually has a character.

Many years ago Lashley pointed out that we are faced with a serious and interesting problem in this kind of phenomenon: most of us have learned to write with the right hand; although none of us who are confirmed right-handers may ever in our lifetime have written with the left hand, we can right now take a pad of paper, close our eyes, put a pencil in the left hand—the nonpreferred hand—and not only write but write upside down. This is remarkably successful usually.

This poses the following problem: on the periphery, which is here tantamount to surface, entirely different muscles are being activated. The order is reversed because you are writing upside down. The direction of the pencil is reversed, and yet quite clearly a pattern is reproduced which you recognize as similar to that produced by the right hand. I think here, analogously, you have to assume that something is stored in the nervous system which corresponds to this patterning. Call it abstract if you wish, because we cannot see it. We cannot cut it out. In animal experiments we have great difficulty interfering by means of a scalpel with this type of neuronal event. These were problems raised by Lashley and confirmed many times since.

CHAIRMAN MILLIKAN: Questions or comments?

DR. THORPE: I would like to ask you questions on one or two points which seem to me to be of great interest. You said, I think, that everyone knows these rules without being taught, without their being formalized or specified. Are you, in fact, saying

that thinking is not simply unvocalized language but that, on the contrary, thinking precedes language? When you refer to the underlying deep structures, would you not say that at least some of these deep structures must be present in animals before the acquisition of any language because they concern the very nature of the perception of or understanding of the world?

There is a good deal of evidence from the work of my colleagues and myself which indicates the highly stimulating effect of mutual practice in promoting normal development towards the specific song pattern in the chaffinch (*Fringilla coelebs*) even though all the birds taking part are equally inexperienced. This suggests very strongly to me that something corresponding to your “deep structures” exists in relation to the development of vocal communication in birds.

DR. CHOMSKY: I think, in fact I am quite convinced, that thinking can proceed perfectly well without language, but I was not presupposing that here. Furthermore, I think one could design an automaton which never thinks but which has the capacity of acquiring any specific human language in a rather short time from a small amount of data.

I would want to put into the automaton an initial specification of what kind of formal object language is, what kind of operations, what kind of mechanisms it has, and leave the automaton just enough leeway so a small amount of data could select one of the possible languages. Having acquired this language, the automaton could produce sentences and assign them semantic interpretation. It would never do anything remotely

like thinking. I think one could describe all these things without ever talking about thinking.

Similarly, I suspect if one knew enough about thinking, one could say many interesting things about it without referring to language. I am quite convinced that thinking is not dependent on language.

I would agree with your statement if you say that thinking precedes language; I am quite sure that is true. I don't doubt that an infant or a dog thinks in some sense and, therefore, that thinking precedes language. I was not presupposing that; in fact, I think the two things are logically, and in large measure factually, independent.

About the question of deep structure, it seems to me, if we look at human beings and ask ourselves the empirical questions as to how they can acquire command of language, that we are led to the postulate, which seems to me fairly well confirmed by evidence, that significant aspects of these deep structures must be present in advance of language acquisition, that there is no inductive operation on language data which leads to the deep structures, but that, rather, the theory of deep structures provides part of a framework for interpreting signals which the organism uses to decide which of the possible languages is the one it is being exposed to.

As far as human beings are concerned, if we simply look at the empirical problem of accounting for the fact that they acquire command of the language, then we are led to the conclusion that they have a certain intrinsic mental structure, and this structure probably specifies the abstract, the general properties of lan-

guage, of deep structures, and of transformations. We can then raise another question, namely, whether this system of mechanisms appears in any other organism, or, for that matter, the question of whether this system of mechanisms plays a role in any other aspect of human intellectual activity, in human cognitive processes, perceptual processes, and so on. I think it would be very surprising if it did not show up in other aspects of human intelligence and, also, somewhere in animal intelligence.

It would be surprising if this complex structure were so specific that it is there just for language and has nothing to do with anything else. I think it would be surprising, but it could be true. I do not think we have any evidence one way or the other.

DR. THORPE: It seemed to me to be rather satisfying to think that these deep structures, or some of them, could be present in animals before language ever evolved; otherwise the evolution of language would be almost impossible. If some of the deep structures can be acquired by normal linguistic experience, then it seems to me that you have got a possibility of language being evolved perhaps a number of times in the course of evolution, without any very great difficulty.

DR. CHOMSKY: I think most of the discussion of evolution of language is complete hand-waving. If one wants to give an account of how some complex structure is evolved, you have to say something more than that a lot of possibilities were tried and this one worked out. If you stop to think about the number of possibilities that can be tried, with no constraints given, then it is quite impossible that anything viable would ever work out.

So, I think if one really wants to give a serious account of the evolutionary process, he must first ask what kind of constraints there are upon the development of complex systems, given certain fixed conditions. Until those constraints are known, I do not see any point in speculation about the evolution of language.

It may turn out, when we learn enough about a complex organism like the brain, that we will discover the only way it can manipulate symbols, when it reaches a certain level of complexity, is this; perhaps this is even a consequence of physical law. If it is a matter of physical law, there is no point speculating about evolution. I am not proposing it as physical law; my point is that at the moment we have as much justification for saying it is physical law that at a certain level of complexity this particular kind of grammatical structure must appear—we have as much reason for saying that as we do for saying that language evolved in some fashion.

One could think of analogous situations in the physical world, where one would draw precisely this kind of conclusion. I am not convinced of the necessity of looking to selectional processes for an explanation of the development of language. If one is looking for an evolutionary explanation of language, I frankly do not see how his problem would be simplified if he were able to discover that other animals have systems analogous to language. I think if he were to discover this he would be faced with a new mystery, namely, how these systems emerged. Adding a new mystery won't help to solve this one. There is a prior question, namely, whether other animals do have systems of this type. Maybe they do, but it does not seem to me that the communication

system of animals, at least, are of this type. If other animals do have systems with the formal properties of human language, I imagine the proper place to look for them is not in the communication systems but, rather, in systems for organization of perception or something of that type.

DR. EVARTS: You spoke of language as being a sound-meaning relationship. You would include, I suppose, sign language in all this. There is nothing special about the sound, if one communicates.

DR. CHOMSKY: I mean "signal." I should have said "signal-meaning correspondence." It is an open question whether the sound part is crucial. It could be but certainly there is little evidence to suggest it is.

DR. SILVERMAN: How can we tap these universals in creating a command of language in deaf children?

DR. CHOMSKY: One would expect that unless the appropriate stimulus conditions are realized, the instinctive behavior would not appear. It may be that the appropriate stimulus condition is hearing enough linguistic noises in your environment. So it just might be that there is no way to tap the system, any more than there is a way of initiating the system of flight in birds without putting them in the situation in which they have to flap their wings.

DR. RICH: Maybe you are not the one I ought to ask this question of. If I understood you correctly, there is a basic set of untransformed frames which the child would have at his command, prelinguistically, or very early. It seems to me at least possible that some transformations are, in a sense, more difficult than others.

Has a study been made of the ages at which the various transformations become available, and could one con-

clude from this anything about the successive psychic difficulty of the transformations?

DR. CHOMSKY: That is a very sensible question, one a lot of people have been worried about for quite a few years. I won't try to summarize the evidence. It seems to me that one general fact about the investigations that have been carried out is that there are two stages that are identifiable. There is a very early stage where in the child's production of speech you do not find many of the mechanisms, and there is a slightly later stage in which you find so many of the mechanisms you cannot begin to describe them, and the transition between these two stages seems fairly rapid.

There has been much careful analysis of language acquisition through its first few months, in which there is a slow increase in one-word sentences, two-word sentences, and three-word sentences, and then the experiment stops. The reason is that at that point development of language increases beyond any possibility of exhaustive description. I am not too convinced about the value of these studies of early stages of language development. These might be like the study of maturation of instinctive capacity before it reaches the appropriate time.

DR. LENNEBERG: I would like to comment on two points and pick up yours first. I am not particularly convinced, after having tried to study the development of language in children, that it is a good question to begin with to ask the time at which suddenly a child begins to use transformations. It makes perfectly good sense to say transformation is a special way of processing material that comes in, transformation of processing on a

chronological level, transformation of processing involved in using words. This can be acceptable, if you think of it in this way. I think even a small child can divide those strings of words that are grammatical from those strings of words that make no sense. Here there is actually involved application of the deep structure that Dr. Chomsky talked about. You apply it and it fits the pattern, and grammatically is the common denominator.

I had originally planned to say that one can show, whenever you have a common denominator of this kind, that you can put this alphabet into this transformation of language where you transform the deep structure to abstract something in the structure into the reality of the sentence. Does this make sense? Let me try to state it another way.

DR. CHOMSKY: I think I agree with that. It seems to me that the deep structures are projected onto the signal by some kind of system of rules which I call transformations. One never produces the deep structures at the earlier stage of language. One does not produce the deep structures which are from the theoretical point of view abstract, and from the neurological point of view probably also deep. But there is some way of mapping these deep structures into signals. The way may differ for the young child, who does not have control of motor mechanisms, and for the child at a later stage.

It is also true, no doubt, that specific aspects of the deep structure and specific projection operations are not available to the child at an early stage. There are some respects in which English is different from German and Chinese, and so on. The child has to learn which of the possible languages is the one he is supposed to be talk-

ing. Before he has learned that, he has, I presume, just some kind of undifferentiated schema which can become Chinese or English or German, when enough data are presented to it.

Nobody knows anything about these mechanisms. It is plausible to assume there is an undifferentiated schema which becomes specified in one way or another and becomes specified and differentiated with more and more complexity and more and more uniqueness through some period of so-called learning. "Learning" is an inappropriate term for that kind of development.

At the latter stage of this process of acquisition, the process of projecting deep structure onto signals will be tremendously more extensive and more complex than at the early stages. There may be much more uniformity in the extensive and complex structure between different individuals than at the very early stages.

I think it is reasonable to suppose that the basic processes of formation of deep structures and projection are being utilized in a very undifferentiated and simple way at the beginning and in a more and more complex way at every stage, until, finally, the normal adult human system is achieved.

DR. CHASE: As I followed the discussion of the deep structure of language, Dr. Chomsky, it seemed quite clear to me that such a structure must exist insofar as we cannot explain the organization of the surface structure of language in its own terms. But it also seemed quite clear that we are not able to specify what the ordering functions are that impose organization on the surface structure of language. Insofar as we cannot specify these ordering operations, and I take it that the word "abstract" is really being

used to say this, then it seems to me that we are not in a position to take a stand on questions of structure and mechanisms, and certainly not on the question of genesis.

It is the last point that I question most seriously, because I think the very best way not to discover the essential substrate in experience for acquiring a capability is to assume that it does not have a substrate in experience. It seems that you are inviting us to consider this possibility with respect to the deep structure of language.

DR. CHOMSKY: I would agree that if we could not specify the deep structures and we could not specify the projection operations, then it would be correct to conclude that we should not speculate about mechanisms and genesis. From what I said, there is no reason why you should guess that we can specify them. In fact, however, I think we can. Of course, anything we can now say is empirical hypothesis, and no doubt will be proved wrong in one respect or another. But I think we can, with a fair degree of confidence, say something rather precise about the exact nature of deep structure and the exact nature of the rules that relate to it. This is the central question of linguistic work. Other linguists would not agree, but this is what I think, at least. It seems to me that a good deal of progress has been made toward giving a precise specification of a universal system of underlying structures and a universal schema which permits some variation from language to language, but not too much, and toward specifying in precise detail the rules of projection, if you like, that map these structures ultimately onto signals.

I have not said anything, really, to justify this, but I think I can refer

you to work which would illustrate these claims. Furthermore, I should mention that there is nothing particularly new about this program or about these conclusions, at least in their general nature. In fact, much of what I have been saying was almost a cliché in the seventeenth and eighteenth centuries, and surprisingly good work was done in what is called universal grammar, a long forgotten tradition, in specifying underlying structures and specifying the abstract operations —by “abstract” I mean “removed from the signal,” not “unspecifiable”—that relate them to surface form. Interestingly, this traditional theory developed out of a framework of rationalist psychology that made assumptions about perception and also assumptions about acquisitions of language that are extremely plausible and closely related to what I think would result from an unbiased investigation of how concepts and systems of this type are acquired. Of course, everything hinges on whether I can substantiate the claim that we can make precise some account of the deep structures and of projection rules.

MRS. RICHARD L. MASLAND: I would like to make a clinical comment about children who are late in developing language and then develop it. Miller has pointed out how important it is to remember the sequence of auditory events. This point has been brought out. For instance, telling the difference between “Bear the eats,” “Eat the bear,” and “The bear eats.” The interesting question is how a child with difficulty in keeping track of the sequence of auditory events can deduce the elements of the deep struc-

tures. This one sees clinically in operation in the child, and one deduces the confusion which may be the basis for the delay. This is in a child with normal hearing acuity.

DR. CHOMSKY: I do not know anything about this.

MRS. MASLAND: Later, when some of the children do acquire language to the point of some fairly normal communication, the semantic elements still become scrambled sometimes. One case in point is a child who said, “My sister bike threw over fence” for “My sister threw her bike over the fence.”

DR. CHOMSKY: Not knowing anything about it, I should not say a word. One can imagine all kinds of possible production defects, filtering defects, let us say, that would lead deep structures to appear not in their normal form but in other forms. For example, consider the phenomenon, which appears apparently with normal children, of producing just major words and not the “grammatical” ones. One might imagine that all that is happening is that there is a filtering device which throws out everything except stressed items. There might be a filtering device or a projection device of some kind that operates on deep structures in some non-normal fashion in the children to give other kinds of signals. One would have to study the material to draw any conclusion.

MRS. MASLAND: We think one of the devices that really can be demonstrated is shortened auditory-memory span.

DR. CHOMSKY: That sounds very plausible.

INTRODUCTION

Some of the outstanding contributions in modern-day psychology serve to "put the train back on the track." David Premack, at least, performs this crucial task by reminding psychologists that whether a particular item is "reinforcing" or "not reinforcing" is a function of the past and present state of the organism. What may reinforce one's behavior today may have no effect tomorrow. The notion seems so forthright that it is sometimes difficult for us to appreciate its news value. Yet the body of American psychology holds that there is a class of items that are "reinforcers" and a class of events that are "reinforced" with no interchange between the two different classes. "Preposterous," says Premack, and in a series of elegant experiments, he shows that even a most universal reinforcer, eating food, can in some instances be the response rewarded by another response such as running! Read carefully and see how the notion opens up vistas and allows you to appreciate the relative nature of the way we assign values to items and events in our life on an ever-changing basis.

REINFORCEMENT THEORY

David Premack

The traditional account of reinforcement provides a nice introduction to the present one, for the latter is a simple transform of the former in the

Reprinted from the Nebraska Symposium on Motivation, 1965, by permission of University of Nebraska Press, Copyright © 1965, pp. 129-48.

following sense. First, the basic assumptions of the traditional account appear to be empirical propositions with a confirmation status so high that further test is unnecessary; actually, the assumptions were little more than common sense, which is to say, they were never tested. Second,

when tested, they proved to be invalid. Thus, all of the major assumptions of the traditional account have essentially the same logical character. They appear to be propositions involving empirical claims, claims so weak as to be essentially trivial. It is only when the propositions are seen to be nonempirical that they become nontrivial.

We may start with the standard definition according to which a reinforcer is any stimulus that, given a certain relation to a response, produces a change in the frequency of that response. Since presumably it is only philosophers who delight in null classes, tradition hastened to add that, of course, there are such stimuli. From there it seemed only reasonable to add: there are other stimuli that do not have this property. The combination, there are some stimuli that have this property and there are others that do not, will be recognized as a familiar version of the first of the several assumptions that make up the traditional descriptive account (e.g., Skinner, 1938, p. 62).

The second empirical claim is that reinforcers are trans-situational (Meehl, 1950). Inevitably, the discovery that any stimulus is effective will be made with some particular base response; the claim of trans-situationality is that a stimulus effective with one response will be effective with all responses (and thus that the discovery could have been equally well made with any response). The classical example is food; it is said to reinforce all responses.

A third empirical claim is that there are two classes of responses, one that is reinforcing but not reinforcible, another that is reinforcible but not reinforcing. This is less an explicit assertion, though versions of it can

be extracted from traditional discussions of goals and drives, than a conclusion that can be reached by examining a sufficient body of experiments. For example, eating and drinking were for many years used only in the role of reinforcer, never as the events that were reinforced; conversely, bar pressing and running were used only as "instrumental activities," never as reinforcers. The segregation here was impressive, there being no instances of crossover between these would-be categories in about forty years of experimentation.

It is important to recognize that the above three assumptions were not viewed as a *theory* of reinforcement, but rather as a first-order, low-level description of the field. Theorists differed as to whether or not reinforcement was coterminous with, say, drive reduction, or whether it was necessary for learning, but they did not differ in their acceptance of the three assumptions.

Are these assumptions, which certainly give the impression of being empirical, actually members of that class in the standard sense of being abstractions from test evidence?

In order for the assertion concerning two classes of stimuli to be an empirical one, the literature must show some body of experiments in which, for any species, some stimuli were shown to change the frequency of a response while other stimuli, given the same temporal relation to the response, failed to produce a change. This kind of functional division of the environment is clearly implied by the classical assumptions, but it is equally clear that no actual division ever took place. That is, it is not possible to list for any species any set of events that have been shown not to be reinforcing. Nor is

this omission the result of being hamstrung by the null hypothesis. The tests were not made because the claim of two classes was never seen to be a claim of any empirical consequence. The fact was too trivial to merit test.

The assumption of trans-situationality cannot be met with the same charge of being wholly nonempirical, for food has been shown to reinforce a number of responses. Food has at one time or another reinforced the bar press, running down an alley, the string pull, and pushing a marble into a hole; doubtless others could be added. But the necessity of piecing together these outcomes from different experiments, which normally would be disallowed, shows that the comparisons are post hoc; the issue was so little seen as empirical that probably not even a master's thesis was devoted to it.

The third assumption could not have been tested (except by accident), for we have already observed that it was tacit, failing even to attain to the status of a necessary condition for being tested, viz., to be made explicit. With a few quite recent exceptions, no attempts were made to reinforce eating or, on the other side of the coin, attempts made to reinforce with the bar press. These events were kept in their proper places. It is no surprise, therefore, that when Meehl's Martian arrived in the animal lab, primed to execute an epistemological reconstruction of reinforcement, he somehow managed to get the responses "right," i.e., managed to reinforce running with eating rather than, as a Martian might, the reverse. (Was the Martian an earthling? He seemed a little too knowing.)

The core assumptions of the traditional account wear empirical masks

which, when they are removed, reveal fairly common-sensical assumptions. The assumptions were not tested; it had already been decided what kinds of events were reinforcing and what kind not. Hence the surprise when lights, sounds, puzzles, etc., proved to be members of the class.

Two choices were given by the "new" reinforcers, either to enlarge the drive category, admit some unforeseen events, and essentially retain the traditional logic of reinforcement, or to drop the standard approach and seek another one.

One reaction to the sudden plethora of reinforcers—lights, sounds, puzzles where a moment ago there had been only food—was to pronounce everything effective, to declare all stimuli reinforcing. But this blanket denial of the former categorical assumption is not one for which any appreciable evidence was ever collected.

The approach we have taken is to observe that reinforcement involves a *relation*, typically between two responses, one that is being reinforced and another that is responsible for the reinforcement. This leads to the following generalization: of any two responses, the more probable response will reinforce the less probable one. Given the generalization, the view of which it is the core can now be set forth in an orderly manner:

1. Anatomically different responses can be compared directly.
2. For any pair of responses, the more probable one will reinforce the less probable one.
3. An indifference principle holds such that the reinforcement value is determined by response probability independent of parameters used to

produce the probability or kind of response that manifests the probability.

4. Reinforcement is a relative property. The most probable response of a set of responses will reinforce all members of the set; the least probable will reinforce no member of the set. However, responses of intermediate probability will reinforce those less probable than themselves but not those more probable than themselves. Intermediate members of the set thus both are and are not reinforcers, depending upon the relative probability of the base response.
5. The reinforcement relation is reversible. If the probability of occurrence of two responses can be reversed in order, so can the reinforcement relation between the two responses.

The logical status of these five assumptions differs from one assumption to the next, and they are presented as a group in the interests of communication rather than for their logical homogeneity. For example, 2 is the central empirical claim, 1 is presupposed by 2, while 4 and 5 are revealing implications of 2.

For the generalization to be tenable, no more than that, it must be possible to compare anatomically different responses. Obviously, it is not possible to talk about the effect that a more probable response may have upon a less probable one if it is not first possible to establish the order of their probabilities. Is it possible to compare anatomically different responses? Tradition may give the impression that it is not, but if so, the impression is misleading, for tradition has not decided against the comparison so

much as it has simply never involved itself in issues that required such comparisons.

Response Methodology

Standard response methodology deals only with what may be called *within-response* comparisons. One and the same response class is compared under different conditions. Approximately 99 per cent of psychological data would appear to be of this sort. As a simple example, consider the classical question of whether rats run more when hungry than when not. Along with more complex issues, this simple question is answered by a within-response comparison. With the use of an activity wheel, a run may be defined as, for example, a 360-degree turn, provided only that a run is counted by this same unit both when the rat is and is not hungry. The use of 360 degrees is, of course, quite arbitrary—359 degrees might be used as well, or even 35 degrees—but the arbitrariness is not injurious to the comparison, for the only criterion in the case of within-response comparison is consistency. That is, whatever the response unit used, it must be the same for the several conditions of the comparison. The criterion for *between-response* comparison is, as we shall see, decidedly stronger, though this is no criticism of standard response procedure for there is no advantage in making criteria stronger than the performance of a task requires.

A further reason for possibly supposing that comparisons between different responses may not be possible is the sheer number of different measures that are used, along with the evident fact of their incommensurability. For example, volume, number of pellets, licks, degrees of turn, num-

ber of bar presses, speed of running are all in use, and quite legitimately, though they are plainly incommensurable. Which of them are to be used for between-response comparison? They cannot all be equally suitable, and yet we appear to have no rule for deciding among them.

In fact, a rule is easily found for while the criterion for within-response comparison is merely consistent application of *any* unit, the criterion for between-response comparison is that *all* responses be measured by the *same* unit. This requirement is not special to responses, of course, but is simply the general requirement for the comparison of any set of items. Thus apples and eggs are commensurable in the case of weight because the same measurement procedure can be applied to both. Similarly, diverse responses will be commensurable, provided they can be measured by the same unit.

A little reflection will show that time, response duration, will fulfill the requirement nicely. Given an operational definition of the onset and offset of a behavior, a clock sensitive to the fulfillment of the operational definition will determine the duration for which the organism is in a state of the designated responding. The clock will not tick more rapidly because the rat is copulating rather than merely eating; the unit of measurement will be invariant from one behavior to the next. Moreover, time is extrinsic to all behaviors and thus can be applied universally, in contrast to all anatomical units, such as the lick, which are intrinsic to given motor sequences and can be used to measure only those sequences composed of the unit in question.

Time not only fulfills the requirement of measurement by a common

unit, thus permitting comparison of anatomically diverse responses, but it takes priority over traditional response criteria such as common-effect-upon-the-environment, or anatomical similarity. In fact, these traditional criteria do not provide sufficient conditions for commensurability, and the attempt to so use them can lead to malpractices. Consider two response classes whose members are either substantially alike anatomically, or which have like effects upon the environment, but which have different average durations. Can the probabilities of the two classes be estimated comparatively from their frequencies? In which case is the organism more apt to be in a state of responding, when it makes 90 one-sec. responses or 10 15-sec. responses? This example alone should make it clear that having a comparable effect upon the environment, such as getting a bar down, is not a sufficient basis for comparing the frequency of two events when the frequency is to be used as an estimate of probability. Comparable average durations of the events in question is at least a necessary condition, and we would argue that it is a sufficient one.

Thus what permits comparing the frequency (number) of anatomically identical responses is no different than what permits the comparison in the case of anatomically disparate responses: average duration of the response events is the same. This point appears to have escaped the attention of several writers (e.g., Hodos & Valenstein, 1962) who have implicitly accepted either same-effect-upon-environment, or anatomical likeness as a sufficient basis for comparing frequencies. The conclusion that in brain stimulation rate will not predict choice would appear to be the result

of this malpractice, and probably can be ratified simply by restricting rate comparisons to events of comparable duration.

Moreover, response duration will provide a strict estimate of probability insofar as the distribution of responding is random throughout the considered interval of time (e.g., Parzan, 1960). Here we may recall the previously noted operating characteristics of the rat, the remarkable extent to which burst duration approximates a random process, which was so surprising. It is always gratifying to put surprises to work.

The above analysis would appear to justify the comparison of anatomically unlike response events, and thus would appear to establish the feasibility of testing the generalization. Tests were directed at four implications: relativity, nonuniqueness of consummatory events, reversibility, and indifference.

Relativity

An original test was made with manipulation responses of Cebus monkeys (Premack, 1963a). By using manipulation—lever pressing of one kind or another—as both reinforcing and reinforced responses, the intent was to show that reinforcement could be produced without the use of goal responses. Tradition had divided the animal's repertoire into goal responses and nongoal responses or "instrumental activities" (the counterpart of neutral stimuli). Accordingly, the standard reinforcement paradigm was: make a goal response contingent upon an instrumental activity. Our gambit in this first test was thus to draw only from the nongoal category, and to demonstrate reinforcement notwithstanding, provided only that one non-

goal response was more probable than the other.

Four manipulanda were used—a lever, plunger, hinged flap or door, and a horizontally operated lever—each of which could be used potentially both as the reinforcing or the reinforced response. A solenoid-operated lock made it possible to render any item inoperable; thus, the animal could be given any pair of items, with one of them locked and its operation made contingent upon the prior operation of the other member of the pair.

A two-stage test procedure was used. First, estimates were obtained of each animal's probability of operating each item. Second, the items were presented in pairs—one free, one locked—with operation of the locked item contingent upon operation of the free item. For example, the door and plunger might be given together, with the plunger free and the door locked. Release of the lock, and thus the opportunity to operate the door, was contingent upon the prior operation of the plunger.

Repeated measures on the items, presented one at a time with no response restrictions, showed the animals to differ substantially. Of four monkeys tested, one showed reliable differences between three of four items, another showed no reliable differences whatever, while two others were intermediate.

In a second procedure, still aimed at estimating response probabilities rather than testing the generalization, the monkeys were presented with *pairs* of items, again with no response restriction. The main purpose was to determine whether simply pairing the items would disturb the ordinal relations among the individual response probabilities. Inevitably, contingency tests will require pairing

items; thus, if merely pairing the items should disturb the rank order of the individual response probabilities, tests of the individual items could not reasonably serve as a basis for predicting reinforcement values. Comparison of the outcome here with that of the previous procedure showed that (1) frequency of responding to any item is apparently always less when the item is paired with another (nonzero) item than when it is presented alone, but (2) the ordinal relations among items are not disturbed simply by the pairing procedure.

Contingencies were next arranged between the pairs of items, using a continuous reinforcement schedule and a rule that required at least one free response between any two contingent ones, thus precluding the hoarding of response possibilities. The six pairs of items (produced by the four manipulanda) generated twelve possible contingency pairs, since in this procedure each item can serve both as a free item on one occasion and as a contingent item on the next occasion. For each pair, a block of four daily contingency sessions were given, followed by a block of four daily extinction sessions, after which the animal was given a new pair of items and the procedure repeated.

The outcome for the four animals differed in a manner generally predictable from their original protocols. Results for Willy, the no-difference animal, were clear cut: rather than showing an increment on any of the five pairs on which he was tested, he declined on all pairs.¹ Willy's overall

decrement would be expected on the following grounds: With no countering incremental effect of reinforcement, the 24-hr. interitem interval that was used in contingency training should produce a lower rate than the 96-hr. interitem interval that was used to obtain the original estimates of response probability (Premack & Bahwell, 1959). The contrast in the rest intervals used for contingency training and for estimation of response probabilities is pertinent to all the results. All increments will be conservative estimates: any increment produced by the contingencies will occur despite an interval that by itself would be predicted to produce a decrement.

The outcome for the two intermediate monkeys, Bimbo and Gimbel, showed one strong and two weak incremental effects; further, in neither *S* did any less probable response increase the frequency of any more probable response. But the clearest predictions possible were those for Chicko,

be low indeed. Would a different set make Willy look like Chicko, or are the differences more a function of the organism than the items? An interesting individual difference would be the tendency to visit a fixed amount of relative preference among the items of any set. Some organisms may display preference among the items of all sets, while others may show no preferences among the items of any set. Thus Willy's flat profile might obtain generality, for all sets, in which case he should be insusceptible to reinforcement. This would be operationally equivalent to a defective reinforcement mechanism, in another idiom to "the salt has lost its savour," or in still a third idiom to the flat-affect model of Meehl's (1962) several models for schizophrenia. We might thus distinguish between two behavioral views of mental disease. In the traditional one, the reinforcement mechanism is intact but the history of reinforcement is untoward, whereas in another, the mechanism itself is defective, either genetically or as a result of a previously untoward reinforcement history.

¹ At least for the present set of items, Willy should be unreinforcible; it may be possible, by habituating him for some items and not for others, to establish the difference needed for reinforcement, but then the ceiling would

who in the first procedure showed three reliably different response probabilities. Indeed, Chicko's protocol made possible three kinds of contingencies: contingent response higher than, less than, and, in one case, about equal to the free response. In brief, the outcome for the three types of contingencies were as follows: (1) contingent response higher than free response produced in all five cases an increment in the free response; (2) contingent less probable than free response produced in all three such cases a decrement in the free response, which is not predicted by the generalization—it says simply that such a case will not produce an increment—but which may be merely the result of the unfavorable intersession interval; and (3) the one case in which the two responses were about equal produced little or no change, increment or decrement.

A last major point of interest in Chicko's data was the evidence concerning the relativity of reinforcement. H was the most probable member of Chicko's response set, P the least, and D intermediate. H reinforced all members of the set; P reinforced none (and instead was itself reinforced by all other members). Considered jointly, the results for H and P appear to substantiate the traditional absolute view: H, which is a reinforcer, reinforces all responses, while P, which is not a reinforcer, reinforces none. This appears to exemplify the "some are, some are not" view, as well as to support the standard trans-situationality assumption—H reinforced not one member of the set but all members. However, the results for D controvert the traditional view; D, which was of intermediate probability, reinforced P but failed to reinforce H. Thus, depending upon

the relative probability of the base response, D both was and was not a reinforcer.

Since the present case appears to represent the first direct test of the trans-situationality assumption, the previous success of the assumption may rest upon a failure to have tested it. Indeed, we are now in a better position to contemplate the success of this assumption. The evidence for it is based entirely upon the demonstration that food (in the starved rat) reinforced all challengers—bar press, string pull, roll a marble. But I think we may see that, rather than proving trans-situationality, this more likely demonstrated that for the given parameters eating was simply the most probable member of the set; certainly all the events reinforced are low operant-level events, while eating in the starved rat can attain substantial probabilities. One response in any set would be predicted to be trans-situational: the most probable member of the set. In using eating and a starved rat, tradition worked exclusively with a most probable member. But a general theory of reinforcement cannot be based upon the special properties of the most probable member of the set. Reinforcement is not confined to this member, and the properties of the other members differ from those of the most probable member. For example, they are not trans-situational.

Reinforcement of Eating and Drinking

Two further basic assumptions that go hand in hand are that consummatory responses are reinforcing and that reinforcement is a reversible relation. Food or water is customarily used to reinforce the bar press or run-

ning, but it is not asked: Can this relation be reversed? Will the bar press or running reinforce eating or drinking?

Are these intervals of time in which eating or drinking are less probable than certain other responses? If there are, then eating and drinking should be reinforcing. The present assumptions do not, of course, make any predictions of the above kind; they deal only with the effect of one response upon another, and thus can make no predictions until after the individual response probabilities are given.²

As to whether or not organisms could be found for which nonconsummatory responses take priority over consummatory, we first considered the need for an advanced organism—something pretty ethereal. The rat did not seem apt to produce so sublime an interval of time. Accordingly, the first test was made with children. Only later did we lift the ban on rats, having found appropriate intervals for them too.

In a first test using children, a pinball machine rewired for continuous operation and a candy dispenser, the two placed side by side, comprised the experimental arrangement (Premack, 1959). Candy consisted of constant-size chocolate bits, delivered one at a time by a conveyer belt into a dish each time the child ate the piece in the dish. Thirty-three children, the entire first-grade class of a public school, served as subjects. Their average age was 6.7 years.

² We have not attempted to provide a general characterization, let alone theory, of events that for a given species lead to a high probability of responding, though such a theory could be of substantial value; e.g., it might be used to deduce the present reinforcement generalization. Cf. Pribram (1963) for an intriguing set of suggestions toward such a theory.

On the first test, both the candy and the pinball machine were available without restriction. Sixty-one per cent of the children made more pinball-machine responses than they ate pieces of candy. (The average duration of the pinball-machine responses was just slightly greater than the visually approximated duration of the ingestion of a piece of candy. Thus the advantage in frequency of pinball responding would convert into an even greater advantage in terms of response duration; since the difference was not great, however, the scores were left in frequency as in the manipulation case.) We thenceforth referred to those who preferred the machine as "manipulators," the others as "eaters."³

On the second test, the availability of the candy and the pinball machine were made subject to either of two contingency relations, E-M or M-E. For E-M, each operation of the machine was contingent upon the prior ingestion of a piece of candy, whereas for M-E, each piece of candy was contingent upon the prior operation of the machine. Both the "manipulators" and the "eaters" were randomly divided, and half of each main group tested under E-M and half under M-E. For a manipulator, E-M was the experimental condition, M-E the control; the reverse held for eaters. What characterizes both experimental conditions is that the more probable response is contingent upon the less

³ "Manipulator" is somewhat of a misnomer. Only some of those who preferred the machine emphasized the manner in which they shot the ball, shooting first from the hip, then with their back turned, all the while doing a kind of soft-shoe routine and practically never glancing at the ball. Others shot the ball with a minimal gesture, then feasted visually upon each new bank of lights the ball lighted up.

probable response (manipulators must eat to manipulate; eaters must manipulate to eat). Control conditions make the more probable response freely available, as in the first test, and thus provide a measure of changes in the less probable response that may occur independent of the contingency.

In both cases, making a more probable response contingent upon a less probable one increased the frequency of the latter. Thus, while candy reinforced manipulation in the eaters, operation of the pinball machine reinforced eating for the manipulators. Increments in the control condition were small compared to those for the experimental condition, and the mean differences within both main groups were significant at less than the 1 percent level.

The results indicate that the summatory response is not unique in being a reinforcer of other responses but not reinforcing itself. Also, there would seem to be little doubt that with sufficient food deprivation, "manipulators" would become "eaters" and thus the response which had been instrumental could be made reinforcing and vice versa.

Reversibility of the Reinforcement Relation

Reversibility of the reinforcement relation was shown ultimately with the rat (Premack, 1962). With free access to both food and an activity wheel but access for only 1 hr./day, mean total drinking time for a group of six female rats was about 240 sec., while mean total running time in the same period was only about 1 min. On the other hand, with free access to both food and water but access to the wheel for only 1 hr./day, mean total drinking time per hour was now only

about 28 sec., while mean total running time in the same period was 329 sec. These are by no means the only parameters that can be used to produce a reversal in the probabilities of drink and run. Combining sucrose concentrations appropriately with force requirements on the wheel will produce the same outcome. So deprivation is hardly the only effective parameter. But whatever parameters are used to reverse the response probabilities, the same prediction follows. It should be possible not only to reinforce drinking with running but also to reverse the reinforcement relation in the same subject simply by changing from one parameter value to another.

The apparatus used to test the predictions consisted of a modified Wahmann activity wheel equipped with a brake and a retractable drinkometer. The rat is placed in the wheel, which itself serves as the test cage. With drinking contingent upon running, the wheel was free but the tube was retracted; a predetermined number of runs brought the tube into the wheel for a predetermined time. Conversely, with running contingent upon drinking, the tube was present at all times but the wheel was locked; a predetermined number of licks released the brake on the wheel for a predetermined time.

The more interesting case, run contingent upon drink, was tested first. Four female albino rats were given daily 1-hr. conditioning sessions, followed by daily 1-hr. extinction and reconditioning sessions. To augment the output, an FR was used in which each five licks freed the wheel for 10 sec. Throughout this training, food and water were continuously available in the home cage; after the last reconditioning session, water was re-

moved from the home cage, and on the following day, training was begun with the reverse contingency—drinking contingent upon running.

With run contingent upon drink, total drink time was increased in all Ss by a factor of from three to five. For operant-level drinking, with only the tube present, mean total drink time was about 28 sec./hr., with both tube and wheel present about 23 sec./hr., and with run contingent upon drink about 98 sec./hr. The first extinction session further increased mean total drink time to about 175 sec./hr., following which drink returned essentially to its base. The picture was readily completed by changing the parameters so as to make drink more probable than run; then, with drink contingent upon run, the conventional increase in run was easily obtained. In summary, parameters were shown that made run more probable than drink and vice versa, and subsequently, it was possible not only to reinforce drink with run but also to reverse the reinforcement relation in the same subjects merely by changing from one set of parameters to the other.

Indifference Principle

Each time a response such as run is successfully predicted to reinforce drink or vice versa, a contribution is made to the indifference principle. This principle holds that prediction of reinforcement can be made without regard to the history of the response probability, for the outcome is indifferent to either the parameters used to produce the probability or the responses that manifest it. Successful prediction for such pairs as run and drink would appear to bear this out, for there is hardly any doubt as to

their being different neurophysiological systems. Indeed, the interest of the indifference principle accrues from this difference; the principle does not deny the neurophysiological difference, which indeed could not be denied, but asserts that despite the difference, the reinforcement value is predictable from a single scale of response probability. The principle says that the neurophysiological difference does not participate in or mediate the reinforcement value, which is determined solely by the duration for which the animal engages in the behavior.

Sharper tests of the point can be made by a somewhat different procedure. With a third response such as the bar press, which is less probable than both drink or run, the increment produced in the bar press by drink and run can be compared. If identical probabilities of run and drink are made contingent upon the bar press, the indifference principle requires that the increment to the bar press be identical. It further requires that if identical probabilities of drink—produced in one case by deprivation, in another by sucrose concentration, in still another by temperature or any other parameter—are made contingent upon the bar press, then, again, the increment to the bar press must be identical. Likewise, it requires that the same experiment in the case of running—where, in principle, identical probabilities of running can be produced by procedures ranging from oestrus to starvation—must also produce the same outcome, identical increments in the bar press.

Since identical response probabilities are easier described than produced, an approximation to this experiment was performed (Premack, 1963b). A set of response probabilities

was produced that included at least one probability of running bounded by probabilities of drinking and at least one probability of drinking bounded by probabilities of running. When subsequently made contingent upon the bar press, the ordered set of run and drink probabilities should produce an identically ordered set of bar-press probabilities.

In Figure 1, bar presses per session are plotted as a function of the associated response probability. The points in the figure are labeled according to the sucrose concentration or force requirement that was used to produce the given drink probability in one case and run probability in the other. The functional relation shown there clearly supports the indif-

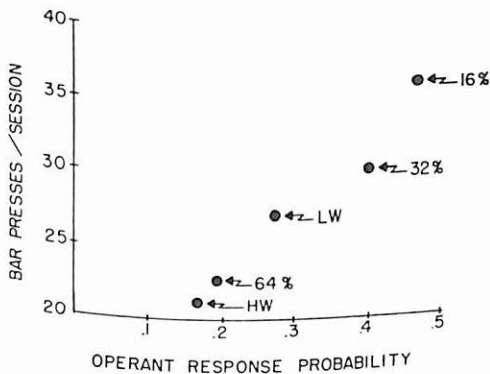


Figure 1

Mean bar presses per session as a function of the associated operant-level response probability. The abscissa shows proportion of the operant-level session for which the animal responded—duration in seconds for which it ran or drank divided by duration of the session. Points are labeled according to the sucrose concentration (16, 32, and 64 per cent) or force requirement (light, LW; heavy, HW) that was used to control the operant-level response probability (after Premack, 1963b).

ference principle: bar pressing increased monotonically with the associated response probability, and did so whether this was a probability of drink or run. These results were more encouraging than others that might be regarded as more dramatic. The universe hardly guarantees that comparable durations of run and drink will have comparable reinforcement effects; yet they clearly appear to.

The indifference principle is an exceptionally strong assumption, not to be established by any one test. If it should prove to be valid for any large number of parameters, it would perhaps contribute more determinately to a unitary-state concept than any traditional drive notion. Are equal probabilities of eating, produced in one case by food deprivation, in another by sucrose concentration, and in still another by lowered temperature, nonetheless the same in their reinforcement value? An animal may be capable of differentiating states of hunger according to the specific antecedents of the hunger; yet it may not differentiate the reinforcement value of the states on any basis other than the associated probabilities of eating.

An Increment Postulate: The Effect of the Base Response

If we apply different more probable responses to a common base response, we may expect the increment to be proportional to the probability of the contingent response. A version of this experiment was shown in the application of varying probabilities of run and drink to the bar press; the greater these probabilities, the greater the increment to the bar press.

But consider the converse experiment. Suppose the same reinforcer is applied to base responses that differ

in their probability of occurrence. Will the reinforcer bring all the base responses to the same level despite their pre-reinforcement differences? Or will the latter be reflected in asymptotic reinforced responding?

A first version of this experiment was done with Chicko via manipulation responding. H, the most probable member of Chicko's set, was applied successively to P, D, and L, responses less probable than H but quite different among themselves. H did not overcome the different operant levels of P, D, and L, but, rather, produced asymptotic rates that were proportional to the original operant levels. Thus, the operant levels showed through the reinforcement.

R. W. Schaeffer performed a more determinate version of this same experiment using groups of rats with drink and run as response materials. The experiment in the rat had been held back by the limited variation in probability of base response that is afforded by the traditional conditioning materials, viz., bar press and eating. But with the reinforcement of drink by run, Schaeffer was given nearly ideal materials. Run was made more probable than drink (mainly by giving only limited access to the wheel), and the probability of drinking was varied widely over the four groups by tube content ranging from nothing to 32 per cent sucrose. After first establishing clear differences in the operant level of drinking, running was made contingent upon drinking in all four groups. Figure 2, which plots asymptotic licking as a function of the pre-reinforcement level of licking, shows that, as in the case of monkey manipulation, the same reinforcer does not bring base responses of different probabilities to the same level. Rather, the asymptotic response level was proportional to

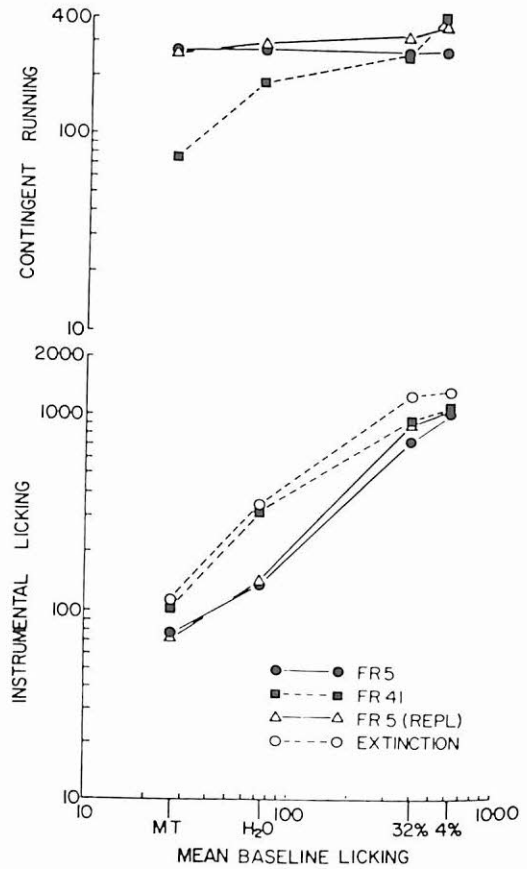


Figure 2

Mean instrumental licks per session as a function of the mean operant-level session immediately preceding the contingency (after Schaeffer, in press).

the pre-reinforcement level; once again, the operant level showed through the reinforcement.

A third version of this experiment has recently been performed by Holstein and Hundt, again with rats, but now using brain-stimulation bar pressing and licking. Drinking a 25 per cent sucrose solution reinforced both brain-stimulation bar pressing and plain bar pressing, but for all rats the duration of bar pressing was greater the greater the pre-reinforcement duration.

INTRODUCTION

In this wide ranging article Tinbergen, the same ethologist who wrote the selection on the bee killers which is reprinted in our Structure of Knowledge section, looks at the fundamental problem of aggression. Tinbergen criticizes, appropriately, the tendency of some people to carry over the results of animal study rather than the methods, for application to humans. The same methods which are useful in studying a variety of lower animals will most likely be useful in studying man; but, it is indeed naive to expect that any specific behavior pattern found in a lower animal will also necessarily be found in man.

Tinbergen's scholarship covers a wide area. Try to understand exactly what he means when he discusses *convergent* and *divergent* evolution. These are important concepts for any student of behavior. As you read this article try to relate it to what you have already read in the "Nature-Nurture" section at the beginning of this collection. It is plainly of the utmost importance that we understand the extent to which internal and external factors influence aggression in man. Tinbergen gives some good suggestions of how we can go about this job and some hints of what we may expect to discover.

ON WAR AND PEACE IN ANIMALS AND MAN

N. Tinbergen

In 1935 Alexis Carrel published a best seller, *Man—The Unknown* (1). Today, more than 30 years later, we biologists have once more the duty to remind our fellowmen that in many respects we are still, to ourselves, unknown. It is true that we now understand a great deal of the way our bodies function. With this understanding came control: medicine.

The ignorance of ourselves which needs to be stressed today is ignorance about our behavior—lack of understanding of the causes and effects of the function of our brains. A scientific understanding of our behavior, leading to its control, may well be the most urgent task that faces mankind today. It is the effects of our behavior that begin to endanger the very survival of our species and, worse, of all life on earth. By our technological achievements we have attained a mastery of our environment that is without precedent in the history of life. But these achievements are rapidly getting out of hand. The consequences of our "rape of the earth" are now assuming critical proportions. With shortsighted recklessness we deplete

the limited natural resources, including even the oxygen and nitrogen of our atmosphere (2). And Rachel Carson's warning (3) is now being followed by those of scientists, who give us an even gloomier picture of the general pollution of air, soil, and water. This pollution is seriously threatening our health and our food supply. Refusal to curb our reproductive behavior has led to the population explosion. And, as if all this were not enough, we are waging war on each other—men are fighting and killing men on a massive scale. It is because the effects of these behavior patterns, and of attitudes that determine our behavior, have now acquired such truly lethal potentialities that I have chosen man's ignorance about his own behavior as the subject of this paper.

I am an ethologist, a zoologist studying animal behavior. What gives a student of animal behavior the temerity to speak about problems of human behavior? Of course the history of medicine provides the answer. We all know that medical research uses animals on a large scale. This makes sense because animals, particularly vertebrates, are, in spite of all differences, so similar to us; they are our blood relations, however distant.

But this use of zoological research for a better understanding of our-

selves is, to most people, acceptable only when we have to do with those bodily functions that we look upon as parts of our physiological machinery—the functions, for instance, of our kidneys, our liver, our hormone-producing glands. The majority of people bridle as soon as it is even suggested that studies of animal behavior could be useful for an understanding, let alone for the control, of our own behavior. They do not want to have their own behavior subjected to scientific scrutiny; they certainly resent being compared with animals, and these rejecting attitudes are both deep-rooted and of complex origin.

But now we are witnessing a turn in this tide of human thought. On the one hand the resistances are weakening, and on the other, a positive awareness is growing of the potentialities of a biology of behavior. This has become quite clear from the great interest aroused by several recent books that are trying, by comparative studies of animals and man, to trace what we could call "the animal roots of human behavior." As examples I select Konrad Lorenz's book *On Aggression* (4) and *The Naked Ape* by Desmond Morris (5). Both books were best sellers from the start. We ethologists are naturally delighted by this sign of rapid growth of interest in our science (even though the growing pains are at times a little hard to endure). But at the same time we are apprehensive, or at least I am.

We are delighted because, from the enormous sales of these and other such books, it is evident that the mental block against self-scrutiny is weakening—that there are masses of people who, so to speak, want to be shaken up.

But I am apprehensive because these books, each admirable in its own

way, are being misread. Very few readers give the authors the benefit of the doubt. Far too many either accept uncritically all that the authors say, or (equally uncritically) reject it all. I believe that this is because both Lorenz and Morris emphasize our knowledge rather than our ignorance (and, in addition, present as knowledge a set of statements which are after all no more than likely guesses). In themselves brilliant, these books could stiffen, at a new level, the attitude of certainty, while what we need is a sense of doubt and wonder, and an urge to investigate, to inquire.

Potential Usefulness of Ethological Studies

Now, in a way, I am going to be just as assertative as Lorenz and Morris, but what I am going to stress is how much we do not know. I shall argue that we shall have to make a major research effort. I am of course fully aware of the fact that much research is already being devoted to problems of human, and even of animal, behavior. I know, for instance, that anthropologists, psychologists, psychiatrists, and others are approaching these problems from many angles. But I shall try to show that the research effort has so far made insufficient use of the potential of ethology. Anthropologists, for instance, are beginning to look at animals, but they restrict their work almost entirely to our nearest relatives, the apes and monkeys. Psychologists do study a larger variety of animals, but even they select mainly higher species. They also ignore certain major problems that we biologists think have to be studied. Psychiatrists, at least many of them, show a disturbing tendency

to apply the *results* rather than the *methods* of ethology to man.

None of these sciences, not even their combined efforts, are as yet parts of one coherent science of behavior. Since behavior is a life process, its study ought to be part of the mainstream of biological research. That is why we zoologists ought to "join the fray." As an ethologist, I am going to try to sketch how my science could assist its sister sciences in their attempts, already well on their way, to make a united, broad-fronted, truly biological attack on the problems of behavior.

I feel that I can cooperate best by discussing what it is in ethology that could be of use to the other behavioral sciences. What we ethologists do not want, what we consider definitely wrong, is uncritical application of our results to man. Instead, I myself at least feel that it is our method of approach, our rationale, that we can offer (6), and also a little simple common sense, and discipline.

The potential usefulness of ethology lies in the fact that, unlike other sciences of behavior, it applies the method or "approach" of biology to the phenomenon behavior. It has developed a set of concepts and terms that allow us to ask:

1. In what ways does this phenomenon (behavior) influence the survival, the success of the animal?
2. What makes behavior happen at any given moment? How does its "machinery" work?
3. How does the behavior machinery develop as the individual grows up?
4. How have the behavior systems of each species evolved until they became what they are now?

The first question, that of survival value, has to do with the effects of behavior; the other three are, each on a different time scale, concerned with its causes.

These four questions are, as many of my fellow biologists will recognize, the major questions that biology has been pursuing for a long time. What ethology is doing could be simply described by saying that, just as biology investigates the functioning of the organs responsible for digestion, respiration, circulation, and so forth, so ethology begins now to do the same with respect to behavior; it investigates the functioning of organs responsible for movement.

I have to make clear that in my opinion it is the comprehensive, integrated attack on all four problems that characterizes ethology. I shall try to show that to ignore the questions of survival value and evolution—as, for instance, most psychologists do—is not only shortsighted but makes it impossible to arrive at an understanding of behavioral problems. Here ethology can make, in fact is already making, positive contributions.

Having stated my case for animal ethology as an essential part of the science of behavior, I will now have to sketch how this could be done. For this I shall have to consider one concrete example, and I select aggression, the most directly lethal of our behaviors. And, for reasons that will become clear, I shall also make a short excursion into problems of education.

Let me first try to define what I mean by aggression. We all understand the term in a vague, general way, but it is, after all, no more than a catchword. In terms of actual behavior, aggression involves approaching an opponent, and, when within reach,

pushing him away, inflicting damage of some kind, or at least forcing stimuli upon him that subdue him. In this description the effect is already implicit: such behavior tends to remove the opponent, or at least to make him change his behavior in such a way that he no longer interferes with the attacker. The methods of attack differ from one species to another, and so do the weapons that are used, the structures that contribute to the effect.

Since I am concentrating on men fighting men, I shall confine myself to intraspecific fighting, and ignore, for instance, fighting between predators and prey. Intraspecific fighting is very common among animals. Many of them fight in two different contexts, which we can call "offensive" and "defensive." Defensive fighting is often shown as a last resort by an animal that, instead of attacking, has been fleeing from an attacker. If it is cornered, it may suddenly turn round upon its enemy and "fight with the courage of despair."

Of the four questions I mentioned before, I shall consider that of the survival value first. Here comparison faces us right at the start with a striking paradox. On the one hand, man is akin to many species of animals in that he fights his own species. But on the other hand he is, among the thousands of species that fight, the only one in which fighting is disruptive.

In animals, intraspecific fighting is usually of distinctive advantage. In addition, all species manage as a rule to settle their disputes without killing one another; in fact, even bloodshed is rare. Man is the only species that is a mass murderer, the only misfit in his own society.

Why should this be so? For an answer, we shall have to turn to the

question of causation: What makes animals and man fight their own species? And why is our species "the odd man out"?

Causation of Aggression

For a fruitful discussion of this question of causation I shall first have to discuss what exactly we mean when we ask it.

I have already indicated that when thinking of causation we have to distinguish between three subquestions, and that these three differ from one another in the stretch of time that is considered. We ask, first: Given an adult animal that fights now and then, what makes each outburst of fighting happen? The time scale in which we consider these recurrent events is usually one of seconds, or minutes. To use an analogy, this subquestion compares with asking what makes a car start or stop each time we use it.

But in asking this same general question of causation ("What makes an animal fight?") we may also be referring to a longer period of time; we may mean "How has the animal, as it grew up, developed this behavior?" This compares roughly with asking how a car has been constructed in the factory. The distinction between these two subquestions remains useful even though we know that many animals continue their development (much slowed down) even after they have attained adulthood. For instance, they may still continue to learn.

Finally, in biology, as in technology, we can extend this time scale even more, and ask: How have the animal species which we observe today—and which we know have evolved from ancestors that were different—how have they acquired their

particular behavior systems during this evolution? Unfortunately, while we know the evolution of cars because they evolved so quickly and have been so fully recorded, the behavior of extinct animals cannot be observed, and has to be reconstructed by indirect methods.

I shall try to justify the claim I made earlier, and show how all these four questions—that of behavior's survival value and the three subquestions of causation—have to enter into the argument if we are to understand the biology of aggression.

Let us first consider the short-term causation: the mechanism of fighting. What makes us fight at any one moment? Lorenz argues in his book that, in animals and in man, there is an internal urge to attack. An individual does not simply wait to be provoked, but, if actual attack has not been possible for some time, this urge to fight builds up until the individual actively seeks the opportunity to indulge in fighting. Aggression, Lorenz claims, can be spontaneous.

But this view has not gone unchallenged. For instance, R. A. Hinde has written a thorough criticism (7), based on recent work on aggression in animals, in which he writes that Lorenz's "arguments for the spontaneity of aggression do not bear examination" and that "the contrary view, expressed in nearly every textbook of comparative psychology . . ." is that fighting "derives principally from the situation"; and even more explicitly: "There is no need to postulate causes that are purely internal to the aggressor" (7, p. 303). At first glance it would seem as if Lorenz and Hinde disagree profoundly. I have read and reread both authors, and it is to me perfectly clear that loose statements and misunderstandings on both sides

have made it appear that there is disagreement where in fact there is something very near to a common opinion. It seems to me that the differences between the two authors lie mainly in the different ways they look at internal and external variables. This in turn seems due to differences of a semantic nature. Lorenz uses the unfortunate term "the spontaneity of aggression." Hinde takes this to mean that external stimuli are in Lorenz's view not necessary at all to make an animal fight. But here he misrepresents Lorenz, for nowhere does Lorenz claim that the internal urge ever makes an animal fight "in vacuo"; somebody or something is attacked. This misunderstanding makes Hinde feel that he has refuted Lorenz's views by saying that "fighting derives principally from the situation." But both authors are fully aware of the fact that fighting is started by a number of variables, of which some are internal and some external. What both authors know, and what cannot be doubted, is that fighting behavior is not like the simple slot machine that produces one platform ticket every time one threepenny bit is inserted. To mention one animal example: a male stickleback does not always show the full fighting behavior in response to an approaching standard opponent; its response varies from none at all to the optimal stimulus on some occasions, to full attack on even a crude dummy at other times. This means that its internal state varies, and in this particular case we know from the work of Hoar (8) that the level of the male sex hormone is an important variable.

Another source of misunderstanding seems to have to do with the stretch of time that the two authors are taking into account. Lorenz un-

doubtedly thinks of the causes of an outburst of fighting in terms of seconds, or hours—perhaps days. Hinde seems to think of events which may have happened further back in time; an event which is at any particular moment “internal” may well in its turn have been influenced previously by external agents. In our stickleback example, the level of male sex hormone is influenced by external agents such as the length of the daily exposure to light over a period of a month or so (9). Or, less far back in time, its readiness to attack may have been influenced by some experience gained, say, half an hour before the fight.

I admit that I have now been spending a great deal of time on what would seem to be a perfectly simple issue: the very first step in the analysis of the short-term causation, which is to distinguish at any given moment between variables within the animal and variables in the environment. It is of course important for our further understanding to unravel the complex interactions between these two worlds, and in particular the physiology of aggressive behavior. A great deal is being discovered about this, but for my present issue there is no use discussing it as long as even the first step in the analysis has not led to a clearly expressed and generally accepted conclusion. We must remember that we are at the moment concerned with the human problem: “What makes men attack each other?” And for this problem the answer to the first stage of our question is of prime importance: Is our readiness to start an attack constant or not? If it were—if our aggressive behavior were the outcome of an apparatus with the properties of the slot machine—all we would have to do would be to control the external situation: to stop providing

threepenny bits. But since our readiness to start an attack is variable, further studies of both the external and the internal variables are vital to such issues as: Can we reduce fighting by lowering the population density, or by withholding provocative stimuli? Can we do so by changing the hormone balance or other physiological variables? Can we perhaps in addition control our development in such a way as to change the dependence on internal and external factors in adult man? However, before discussing development, I must first return to the fact that I have mentioned before, namely, that man is, among the thousands of other species that fight, the only mass murderer. How do animals in their intraspecific disputes avoid bloodshed?

The Importance of “Fear”

The clue to this problem is to recognize the simple fact that aggression in animals rarely occurs in pure form; it is only one of two components of an adaptive system. This is most clearly seen in territorial behavior, although it is also true of most other types of hostile behavior. Members of territorial species divide, among themselves, the available living space and opportunities by each individual defending its home range against competitors. Now in this system of parceling our living space, avoidance plays as important a part as attack. Put very briefly, animals of territorial species, once they have settled on a territory, attack intruders, but an animal that is still searching for a suitable territory or finds itself outside its home range withdraws when it meets with an already established owner. In terms of function, once you have taken possession of a territory, it pays to drive

off competitors; but when you are still looking for a territory (or meet your neighbor at your common boundary), your chances of success are improved by avoiding such established owners. The ruthless fighter who "knows no fear" does not get very far. For an understanding of what follows, this fact, that hostile clashes are controlled by what we could call the "attack-avoidance system," is essential.

When neighboring territory owners meet near their common boundary, both attack behavior and withdrawal behavior are elicited in both animals; each of the two is in a state of motivational conflict. We know a great deal about the variety of movements that appear when these two conflicting, incompatible behaviors are elicited. Many of these expressions of a motivational conflict have, in the course of evolution, acquired signal function; in colloquial language, they signal "Keep out!" We deduce this from the fact that opponents respond to them in an appropriate way: instead of proceeding to intrude, which would require the use of force, trespassers withdraw, and neighbors are contained by each other. This is how such animals have managed to have all the advantages of their hostile behavior without the disadvantages: they divide their living space in a bloodless way by using as distance-keeping devices these conflict movements ("threat") rather than actual fighting.

Group Territories

In order to see our wars in their correct biological perspective one more comparison with animals is useful. So far I have discussed animal species that defend individual or at best pair territories. But there are also

animals which possess and defend territories belonging to a group, or a clan (10).

Now it is an essential aspect of group territorialism that the members of a group unite when in hostile confrontation with another group that approaches, or crosses into their feeding territory. The uniting and the aggression are equally important. It is essential to realize that group territorialism does not exclude hostile relations on lower levels when the group is on its own. For instance, within a group there is often a peck order. And within the group there may be individual or pair territories. But frictions due to these relationships fade away during a clash between groups. This temporary elimination is done by means of so-called appeasement and reassurance signals. They indicate "I am a friend," and so diminish the risk that, in the general flare-up of anger, any animal "takes it out" on a fellow member of the same group (11). Clans meet clans as units, and each individual in an intergroup clash, while united with its fellow-members, is (as in interindividual clashes) torn between attack and withdrawal, and postures and shouts rather than attacks.

We must now examine the hypothesis (which I consider the most likely one) that man still carries with him the animal heritage of group territoriality. This is a question concerning man's evolutionary origin, and here we are, by the very nature of the subject, forced to speculate. Because I am going to say something about the behavior of our ancestors of, say, 100,000 years ago, I have to discuss briefly a matter of methodology. It is known to all biologists (but unfortunately unknown to most psychologists) that comparison of present-

day species can give us a deep insight, with a probability closely approaching certainty, into the evolutionary history of animal species. Even where fossil evidence is lacking, this comparative method alone can do this. It has to be stressed that this comparison is a highly sophisticated method, and not merely a matter of saying that species A is different from species B (12). The basic procedure is this. We interpret differences between really allied species as the result of adaptive divergent evolution from common stock, and we interpret similarities between nonallied species as adaptive convergencies to similar ways of life. By studying the adaptive functions of species characteristics we understand how natural selection can have produced both these divergencies and convergencies. To mention one striking example: even if we had no fossil evidence, we could, by this method alone, recognize whales for what they are—mammals that have returned to the water, and, in doing so, have developed some similarities to fish. This special type of comparison, which has been applied so successfully by students of the structure of animals, has now also been used, and with equal success, in several studies of animal behavior. Two approaches have been applied. One is to see in what respects species of very different origin have convergently adapted to a similar way of life. Von Haartman (13) has applied this to a study of birds of many types that nest in holes—an anti-predator safety device. All such hole-nesters center their territorial fighting on a suitable nest hole. Their courtship consists of luring a female to this hole (often with the use of bright color patterns). Their young gape when a general darkening signals the arrival of the

parent. All but the most recently adapted species lay uniformly colored, white or light blue eggs that can easily be seen by the parent.

An example of adaptive divergence has been studied by Cullen (14). Among all the gulls, the kittiwake is unique in that it nests on very narrow ledges on sheer cliffs. Over 20 peculiarities of this species have been recognized by Mrs. Cullen as vital adaptations to this particular habitat.

These and several similar studies (15) demonstrate how comparison reveals, in each species, systems of interrelated, and very intricate adaptive features. In this work, speculation is now being followed by careful experimental checking. It would be tempting to elaborate on this, but I must return to our own unfortunate species.

Now, when we include the "Naked Ape" in our comparative studies, it becomes likely (as has been recently worked out in great detail by Morris) that man is a "social Ape who has turned carnivore" (16). On the one hand he is a social primate; on the other, he has developed similarities to wolves, lions and hyenas. In our present context one thing seems to stand out clearly, a conclusion that seems to me of paramount importance to all of us, and yet has not yet been fully accepted as such. As a social, hunting primate, man must originally have been organized on the principle of group territories.

Ethologists tend to believe that we still carry with us a number of behavioral characteristics of our animal ancestors, which cannot be eliminated by different ways of upbringing, and that our group territorialism is one of those ancestral characters. I shall discuss the problem of the modifiability of our behavior later, but it is use-

ful to point out here that even if our behavior were much more modifiable than Lorenz maintains, our cultural evolution, which resulted in the parceling-out of our living space on lines of tribal, national, and now even "bloc" areas, would, if anything, have tended to enhance group territorialism.

Group Territorialism in Man?

I put so much emphasis on this issue of group territorialism because most writers who have tried to apply ethology to man have done this in the wrong way. They have made the mistake, to which I objected before, of uncritically extrapolating the results of animal studies to man. They try to explain man's behavior by using facts that are valid only of some of the animals we studied. And, as ethologists keep stressing, no two species behave alike. Therefore, instead of taking this easy way out, we ought to study man in his own right. And I repeat that the message of the ethologists is that the methods, rather than the results, of ethology should be used for such a study.

Now, the notion of territory was developed by zoologists (to be precise, by ornithologists, 17), and because individual and pair territories are found in so many more species than group territories (which are particularly rare among birds), most animal studies were concerned with such individual and pair territories. Now such low-level territories do occur in man, as does another form of hostile behavior, the peck order. But the problems created by such low-level frictions are not serious; they can, within a community, be kept in check by the apparatus of law and order; peace within national boundaries can be en-

forced. In order to understand what makes us go to war, we have to recognize that man behaves very much like a group-territorial species. We too unite in the face of an outside danger to the group; we "forget our differences." We too have threat gestures, for instance, angry facial expressions. And all of us use reassurance and appeasement signals, such as a friendly smile. And (unlike speech) these are universally understood; they are cross-cultural; they are species-specific. And, incidentally, even within a group sharing a common language, they are often more reliable guides to a man's intentions than speech, for speech (as we know now) rarely reflects our true motives, but our facial expressions often "give us away."

If I may digress for a moment: it is humiliating to us ethologists that many nonscientists, particularly novelists and actors, intuitively understand our sign language much better than we scientists ourselves do. Worse, there is a category of human beings who understand intuitively more about the causation of our aggressive behavior: the great demagogues. They have applied this knowledge in order to control our behavior in the most clever ways, and often for the most evil purposes. For instance, Hitler (who had modern mass communication at his disposal, which allowed him to inflame a whole nation) played on both fighting tendencies. The "defensive" fighting was whipped up by his passionate statements about "living space," "encirclement," Jewry, and Freemasonry as threatening powers which made the Germans feel "cornered." The "attack fighting" was similarly set ablaze by playing the myth of the *Herrenvolk*. We must make sure that mankind has learned its lesson and will never forget how

disastrous the joint effects have been—if only one of the major nations were led now by a man like Hitler, life on earth would be wiped out.

I have argued my case for concentrating on studies of group territoriality rather than on other types of aggression. I must now return, in this context, to the problem of man the mass murderer. Why don't we settle even our international disputes by the relatively harmless, animal method of threat? Why have we become unhinged so that so often our attack erupts without being kept in check by fear? It is not that we have no fear, nor that we have no other inhibitions against killing. This problem has to be considered first of all in the general context of the consequences of man having embarked on a new type of evolution.

Cultural Evolution

Man has the ability, unparalleled in scale in the animal kingdom, of passing on his experiences from one generation to the next. By this accumulative and exponentially growing process, which we call cultural evolution, he has been able to change his environment progressively out of all recognition. And this includes the social environment. This new type of evolution proceeds at an incomparably faster pace than genetic evolution. Genetically we have not evolved very strikingly since Cro-Magnon man, but culturally we have changed beyond recognition, and are changing at an ever-increasing rate. It is of course true that we are highly adjustable individually, and so could hope to keep pace with these changes. But I am not alone in believing that this behavioral adjustability, like all types of modifiability, has its limits. These

limits are imposed upon us by our hereditary constitution, a constitution which can only change with the far slower speed of genetic evolution. There are good grounds for the conclusion that man's limited behavioral adjustability has been outpaced by the culturally determined changes in his social environment, and that this is why man is now a misfit in his own society.

We can now, at last, return to the problem of war, of uninhibited mass killing. It seems quite clear that our cultural evolution is at the root of the trouble. It is our cultural evolution that has caused the population explosion. In a nutshell, medical science, aiming at the reduction of suffering, has, in doing so, prolonged life for many individuals as well—prolonged it to well beyond the point at which they produce offspring. Unlike the situation in any wild species, recruitment to the human population consistently surpasses losses through mortality. Agricultural and technical know-how have enabled us to grow food and to exploit other natural resources to such an extent that we can still feed (though only just) the enormous numbers of human beings on our crowded planet. The result is that we now live at a far higher density than that in which genetic evolution has molded our species. This, together with long-distance communication, leads to far more frequent, in fact to continuous, intergroup contacts, and so to continuous external provocation of aggression. Yet this alone would not explain our increased tendency to kill each other; it would merely lead to continuous threat behavior.

The upsetting of the balance between aggression and fear (and this is what causes war) is due to at least

three other consequences of cultural evolution. It is an old cultural phenomenon that warriors are both brainwashed and bullied into all-out fighting. They are brainwashed into believing that fleeing—originally, as we have seen, an adaptive type of behavior—is despicable, “cowardly.” This seems to me due to the fact that man, accepting that in moral issues death might be preferable to fleeing, has falsely applied the moral concept of “cowardice” to matters of mere practical importance—to the dividing of living space. The fact that our soldiers are also bullied into all-out fighting (by penalizing fleeing in battle) is too well known to deserve elaboration.

Another cultural excess is our ability to make and use killing tools, especially long-range weapons. These make killing easy, not only because a spear or a club inflicts, with the same effort, so much more damage than a fist, but also, and mainly, because the use of long-range weapons prevents the victim from reaching his attacker with his appeasement, reassurance, and distress signals. Very few aircrews who are willing, indeed eager, to drop their bombs “on target” would be willing to strangle, stab, or burn children (or, for that matter, adults) with their own hands; they would stop short of killing, in response to the appeasement and distress signals of their opponents.

These three factors alone would be sufficient to explain how we have become such unhinged killers. But I have to stress once more that all this, however convincing it may seem, must still be studied more thoroughly.

There is a frightening, and ironical paradox in this conclusion: that the human brain, the finest life-preserving device created by evolution, has

made our species so successful in mastering the outside world that it suddenly finds itself taken off guard. One could say that our cortex and our brainstem (our “reason” and our “instincts”) are at loggerheads. Together they have created a new social environment in which, rather than ensuring our survival, they are about to do the opposite. The brain finds itself seriously threatened by an enemy of its own making. It is its own enemy. We simply have to understand this enemy.

The Development of Behavior

I must now leave the question of the moment-to-moment control of fighting, and, looking further back in time, turn to the development of aggressive behavior in the growing individual. Again we will start from the human problem. This, in the present context, is whether it is within our power to control development in such a way that we reduce or eliminate fighting among adults. Can or cannot education in the widest sense produce nonaggressive men?

The first step in the consideration of this problem is again to distinguish between external and internal influences, but now we must apply this to the growth, the changing, of the behavioral machinery during the individual's development. Here again the way in which we phrase our questions and our conclusions is of the utmost importance.

In order to discuss this issue fruitfully, I have to start once more by considering it in a wider context, which is now that of the “nature-nurture” problem with respect to behavior in general. This has been discussed more fully by Lorenz in his book *Evolution and Modification of*

Behaviour (18); for a discussion of the environmentalist point of view I refer to the various works of Schneirla (see 19).

Lorenz tends to classify behavior types into innate and acquired or learned behavior. Schneirla rejects this dichotomy into two classes of behavior. He stresses that the developmental process, of behavior as well as of other functions, should be considered, and also that this development forms a highly complicated series of interactions between the growing organism and its environment. I have gradually become convinced that the clue to this difference in approach is to be found in a difference in aims between the two authors. Lorenz claims that "we are justified in leaving, at least for the time being, to the care of the experimental embryologists all those questions which are concerned with the chains of physiological causation leading from the genome to the development of . . . neurosensory structures" (18, p. 43). In other words, he deliberately refrains from starting his analysis of development prior to the stage at which a fully coordinated behavior is performed for the first time. If one in this way restricts one's studies to the later stages of development, then a classification in "innate" and "learned" behavior, or behavior components, can be considered quite justified. And there was a time, some 30 years ago, when the almost grotesquely environmentalist bias of psychology made it imperative for ethologists to stress the extent to which behavior patterns could appear in perfect or near-perfect form without the aid of anything that could be properly called learning. But I now agree (however belatedly) with Schneirla that we must extend our

interest to earlier stages of development and embark on a full program of experimental embryology of behavior. When we do this, we discover that interactions with the environment can indeed occur at early stages. These interactions may concern small components of the total machinery of a fully functional behavior pattern, and many of them cannot possibly be called learning. But they are interactions with the environment, and must be taken into account if we follow in the footsteps of the experimental embryologists, and extend our field of interest to the entire sequence of events which lead from the blueprints contained in the zygote to the fully functioning, behaving animal. We simply have to do this if we want an answer to the question to what extent the development of behavior can be influenced from the outside.

When we follow this procedure the rigid distinction between "innate" or unmodifiable and "acquired" or modifiable behavior patterns becomes far less sharp. This is owing to the discovery, on the one hand, that "innate" patterns may contain elements that at an early stage developed in interaction with the environment, and, on the other hand, that learning is, from step to step, limited by internally imposed restrictions.

To illustrate the first point, I take the development of the sensory cells in the retina of the eye. Knoll has shown (20) that the rods in the eyes of tadpoles cannot function properly unless they have first been exposed to light. This means that, although any visually guided response of a tadpole may well, in its integrated form, be "innate" in Lorenz's sense, it is so only in the sense of "nonlearned," not in that of "having grown without interaction with the environment."

Now it has been shown by Cullen (21) that male sticklebacks reared from the egg in complete isolation from other animals will, when adult, show full fighting behavior to other males and courtship behavior to females when faced with them for the first time in their lives. This is admittedly an important fact, demonstrating that the various recognized forms of learning do not enter into the programming of these integrated patterns. This is a demonstration of what Lorenz calls an "innate response." But it does not exclude the possibility that parts of the machinery so employed may, at an earlier stage, have been influenced by the environment, as in the case of the tadpoles.

Second, there are also behavior patterns which do appear in the inexperienced animal, but in an incomplete form, and which require additional development through learning. Thorpe has analyzed a clear example of this: when young male chaffinches reared alone begin to produce their song for the first time, they utter a very imperfect warble; this develops into the full song only if, at a certain sensitive stage, the young birds have heard the full song of an adult male (22).

By far the most interesting aspect of such intermediates between innate and acquired behavior is the fact that learning is not indiscriminate, but is guided by a certain selectiveness on the part of the animal. This fact has been dimly recognized long ago; the early ethologists have often pointed out that different, even closely related, species learn different things even when developing the same behavior patterns. This has been emphasized by Lorenz's use of the term "innate teaching mechanism." Other authors

use the word "template" in the same context. The best example I know is once more taken from the development of song in certain birds. As I have mentioned, the males of some birds acquire their full song by changing their basic repertoire to resemble the song of adults, which they have to hear during a special sensitive period some months before they sing themselves. It is in this sensitive period that they acquire, without as yet producing the song, the knowledge of "what the song ought to be like." In technical terms, the bird formed a *Sollwert* (23) (literally, "should-value," an ideal) for the feedback they receive when they hear their own first attempts. Experiments have shown (24) that such birds, when they start to sing, do three things: they listen to what they produce; they notice the difference between this feedback and the ideal song; and they correct their next performance.

This example, while demonstrating an internal teaching mechanism, shows, at the same time, that Lorenz made his concept too narrow when he coined the term "innate teaching mechanism." The birds have developed a teaching mechanism, but while it is true that it is internal, it is not innate; the birds have acquired it by listening to their father's song.

These examples show that if behavior studies are to catch up with experimental embryology our aims, our concepts, and our terms must be continually revised.

Before returning to aggression, I should like to elaborate a little further on general aspects of behavior development, because this will enable me to show the value of animal studies in another context, that of education.

Comparative studies, of different

animal species, of different behavior patterns, and of different stages of development, begin to suggest that wherever learning takes a hand in development, it is guided by such *Sollwerte* or templates for the proper feedback, the feedback that reinforces. And it becomes clear that these various *Sollwerte* are of a bewildering variety. In human education one aspect of this has been emphasized in particular, and even applied in the use of teaching machines: the requirement that the reward, in order to have maximum effect, must be immediate. Skinner has stressed this so much because in our own teaching we have imposed an unnatural delay between, say, taking in homework, and giving the pupil his reward in the form of a mark. But we can learn more from animal studies than the need for immediacy of reward. The type of reward is also of great importance, and this may vary from task to task, from stage to stage, from occasion to occasion; the awards may be of almost infinite variety.

Here I have to discuss briefly a behavior of which I have so far been unable to find the equivalent in the development of structure. This is exploratory behavior. By this we mean a kind of behavior in which the animal sets out to acquire as much information about an object or a situation as it can possibly get. The behavior is intricately adapted to this end, and it terminates when the information has been stored, when the animal has incorporated it in its learned knowledge. This exploration (subjectively we speak of "curiosity") is not confined to the acquisition of information about the external world alone; at least mammals explore their own movements a great deal, and in this

way "master new skills." Again, in this exploratory behavior, *Sollwerte* of expected, "hoped-for" feedbacks play their part.

Without going into more detail, we can characterize the picture we begin to get of the development of behavior as a series, or rather a web, of events, starting with innate programming instructions contained in the zygote, which straightaway begin to interact with the environment; this interaction may be discontinuous, in that periods of predominantly internal development alternate with periods of interaction, or sensitive periods. The interaction is enhanced by active exploration; it is steered by selective *Sollwerte* of great variety; and stage by stage this process ramifies; level upon level of ever-increasing complexity is being incorporated into the programming.

Apply what we have heard for a moment to playing children (I do not, of course, distinguish sharply between "play" and "learning"). At a certain age a child begins to use, say, building blocks. It will at first manipulate them in various ways, one at a time. Each way of manipulating acts as exploratory behavior: the child learns what a block looks, feels, tastes like, and so forth, and also how to put it down so that it stands stably.

Each of these stages "peters out" when the child knows *what it wanted to find out*. But as the development proceeds, a new level of exploration is added: the child discovers that it can put one block on top of the other; it constructs. The new discovery leads to repetition and variation, for each child develops, at some stage, a desire and a set of *Sollwerte* for such effects of construction, and acts out to the full this new level of exploratory be-

havior. In addition, already at this stage the *Sollwert* or ideal does not merely contain what the blocks do, but also what, for instance, the mother does; her approval, her shared enjoyment, is also of great importance. Just as an exploring animal, the child builds a kind of inverted pyramid of experience, built of layers, each set off by a new wave of exploration and each directed by new sets of *Sollwerte*, and so its development "snowballs." All these phases may well have more or less limited sensitive periods, which determine when the fullest effect can be obtained, and when the child is ready for the next step. More important still, if the opportunity for the next stage is offered either too early or too late, development may be damaged, including the development of motivational and emotional attitudes.

Of course gifted teachers of many generations have known all these things (25) or some of them, but the glimpses of insight have not been fully and scientifically systematized. In human education, this would of course involve experimentation. This need not worry us too much, because in our search for better educational procedures we are in effect experimenting on our children all the time. Also, children are fortunately incredibly resilient, and most grow up into pretty viable adults in spite of our fumbling educational efforts. Yet there is, of course, a limit to what we will allow ourselves, and this, I should like to emphasize, is where animal studies may well become even more important than they are already.

Can Education End Aggression?

Returning now to the development of animal and human aggression, I

hope to have made at least several things clear: that behavior development is a very complex phenomenon indeed; that we have only begun to analyze it in animals; that with respect to man we are, if anything, behind in comparison with animal studies; and that I cannot do otherwise than repeat what I said in the beginning: we must make a major research effort. In this effort animal studies can help, but we are still very far from drawing very definite conclusions with regard to our question: To what extent shall we be able to render man less aggressive through manipulation of the environment, that is, by educational measures?

In such a situation personal opinions naturally vary a great deal. I do not hesitate to give as my personal opinion that Lorenz's book *On Aggression*, in spite of its assertativeness, in spite of factual mistakes, and in spite of the many possibilities of misunderstandings that are due to the lack of a common language among students of behavior—that this work must be taken more seriously as a positive contribution to our problem than many critics have done. Lorenz is, in my opinion, right in claiming that elimination, through education, of the internal urge to fight will turn out to be very difficult, if not impossible.

Everything I have said so far seems to me to allow for only one conclusion. Apart from doing our utmost to return to a reasonable population density, apart from stopping the progressive depletion and pollution of our habitat, we must pursue the biological study of animal behavior for clarifying problems of human behavior of such magnitude as that of our aggression, and of education.

But research takes a long time, and

we must remember that there are experts who forecast worldwide famine 10 to 20 years from now; and that we have enough weapons to wipe out all human life on earth. Whatever the causation of our aggression, the simple fact is that for the time being we are saddled with it. This means that there is a crying need for a crash program, for finding ways and means for keeping our intergroup aggression in check. This is of course in practice infinitely more difficult than controlling our intranational frictions; we have as yet not got a truly international police force. But there is hope for avoiding all-out war because, for the first time in history, we are afraid of killing ourselves by the lethal radiation effects even of bombs that we could drop in the enemy's territory. Our politicians know this. And as long as there is this hope, there is every reason to try and learn what we can from animal studies. Here again they can be of help. We have already seen that animal opponents meeting in a hostile clash avoid bloodshed by using the expressions of their motivational conflicts as intimidating signals. Ethologists have studied such conflict movements in some detail (26), and have found that they are of a variety of types. The most instructive of these is the redirected attack; instead of attacking the provoking, yet dreaded, opponent, animals often attack something else, often even an inanimate object. We ourselves bang the table with our fists. Redirection includes something like sublimation, a term attaching a value judgment to the redirection. As a species with group territories, humans, like hyenas, unite when meeting a common enemy. We do already sublimate our group aggression. The Dutch feel united in their fight

against the sea. Scientists do attack their problems together. The space program—surely a mainly military effort—is an up-to-date example. I would not like to claim, as Lorenz does, that redirected attack exhausts the aggressive urge. We know from soccer matches and from animal work how aggressive behavior has two simultaneous, but opposite effects: a waning effect, and one of self-inflammation, of mass hysteria, such as recently seen in Cairo. Of these two the inflammatory effect often wins. But if aggression were used successfully as the motive force behind nonkilling and even useful activities, self-stimulation need not be a danger; in our short-term cure we are not aiming at the elimination of aggressiveness, but at "taking the sting out of it."

Of all sublimated activities, scientific research would seem to offer the best opportunities for deflecting and sublimating our aggression. And, once we recognize that it is the disrupted relation between our own behavior and our environment that forms our most deadly enemy, what could be better than uniting, at the front or behind the lines, in the scientific attack on our own behavioral problems?

I stress "behind the lines." The whole population should be made to feel that it participates in the struggle. This is why scientists will always have the duty to inform their fellow-men of what they are doing, of the relevance and the importance of their work. And this is not only a duty, it can give intense satisfaction.

I have come full circle. For both the long-term and the short-term remedies at least we scientists will have to sublimate our aggression into an all-out attack on the enemy within. For this the enemy must be recognized for

what it is: our unknown selves, or, deeper down, our refusal to admit that man is, to himself, unknown.

I should like to conclude by saying a few words to my colleagues of the younger generation. Of course we all hope that, by muddling along until we have acquired better understanding, self-annihilation either by the "whimper of famine" or by the "bang of war" can be avoided. For this, we must on the one hand trust, on the other help (and urge) our politicians. But it is no use denying that the chances of designing the necessary preventive measures are small, let alone the chances of carrying them out. Even birth control still offers a major problem.

It is difficult for my generation to know how seriously you take the danger of mankind destroying his own species. But those who share the apprehension of my generation might perhaps, with us, derive strength from keeping alive the thought that has helped so many of us in the past when faced with the possibility of imminent death. Scientific research is one of the finest occupations of our mind. It is, with art and religion, one of the uniquely human ways of meeting nature, in fact, the most active way. If we are to succumb, and even if this were to be ultimately due to our own stupidity, we could still, so to speak, redeem our species. We could at least go down with some dignity, by using our brain for one of its supreme tasks, by exploring to the end.

References

1. CARREL, A., *L'Homme, cet Inconnu* (Librairie Plon, Paris, 1935).
2. AAAS Annual Meeting, 1967 [See *New Scientist* 37, 5 (1968)].
3. CARSON, R., *Silent Spring* (Houghton Mifflin, Boston, 1962).
4. LORENZ, K., *On Aggression* (Methuen, London, 1966).
5. MORRIS, D., *The Naked Ape* (Jonathan Cape, London, 1967).
6. TINBERGEN, N., *Z. Tierpsychol.* 20, 410 (1954).
7. HINDE, R. A., *New Society* 9, 302 (1967).
8. HOAR, W. S., *Animal Behaviour* 10, 247 (1962).
9. BAGGERMAN, B., in *Symp. Soc. Exp. Biol.* 20, 427 (1965).
10. KRUUK, H., *New Scientist* 30, 849 (1966).
11. TINBERGEN, N., *Z. Tierpsychol.* 16, 651 (1959); *Zool. Mededelingen* 39, 209 (1964).
12. ———, *Behaviour* 15, 1-70 (1959).
13. VON HAARTMAN, L., *Evolution* 11, 339 (1957).
14. CULLEN, E., *Ibid.* 99, 275 (1957).
15. CROOK, J. H., *Symp. Zool. Soc. London* 14, 181 (1965).
16. FREEMAN, D., *Inst. Biol. Symp.* 13, 109 (1964); MORRIS, D., Ed., *Primate Ethology* (Weidenfeld and Nicolson, London, 1967).
17. HOWARD, H. E., *Territory in Bird Life* (Murray, London, 1920); Hinde, R. A., et al., *Ibis* 98, 340-530 (1956).
18. LORENZ, K., *Evolution and Modification of Behaviour* (Methuen, London, 1966).
19. SCHNEIRLA, T. C., *Quart. Rev. Biol.* 41, 283 (1966).
20. KNOLL, M. D., *Z. Vergleich. Physiol.* 38, 219 (1956).
21. CULLEN, E., *Final Rept. Contr. AF 61 (052)-29*, USAFRDC, 1-23 (1961).
22. THORPE, W. H., *Bird-Song* (Cambridge Univ. Press, New York, 1961).

23. VON HOLST, E. and H. MITTELSTAEDT, *Naturwissenschaften* 37, 464 (1950).
24. KONISHI, M., *Z. Tierpsychol.* 22, 770 (1965); NOTTEBOHM, F., *Proc. 14th Intern. Ornithol. Congr.* 265-280 (1967).
25. STANDING, E. M., *Maria Montessori* (New American Library, New York, 1962).
26. TINBERGEN, N., in *The Pathology and Treatment of Sexual Deviation*, I. Rosen, Ed. (Oxford Univ. Press, London, 1964), pp. 3-23; JONES, N. B., *Wildfowl Trust 11th Ann. Rept.*, 46-52 (1960); SEVENSTER, P., *Behaviour, Suppl.* 9, 1-170 (1961); ROWELL, F., *Animal Behaviour* 9, 38 (1961).

INTRODUCTION

Hypnotism has long been a fascinating subject. In the following selection from his book *Principles of Psychology*, William James recounts much that was known in his time about this still-mysterious process.

HYPNOTISM

William James

Modes of Operating, and Susceptibility

The 'hypnotic,' 'mesmeric,' or 'magnetic' trance *can be induced in various ways*, each operator having his pet method. The simplest one is to leave the subject seated by himself, telling him that if he close his eyes and relax his muscles and, as far as possible, think of vacancy, in a few minutes he will 'go off.' On returning in ten minutes you may find him effectually hypnotized. Braid used to make his subjects look at a bright button held near their forehead until their eyes spontaneously closed. The older mesmerists made 'passes' in a

downward direction over the face and body, but without contact. Stroking the skin of the head, face, arms and hands, especially that of the region round the brows and eyes, will have the same effect. Staring into the eyes of the subject until the latter droop; making him listen to a watch's ticking; or simply making him close his eyes for a minute whilst you describe to him the feeling of falling into sleep, 'talk sleep' to him, are equally efficacious methods in the hands of some operators; whilst with trained subjects any method whatever from which they have been led by previous suggestion to expect results will be successful.¹ The touching of an object

Reprinted with permission of the publisher from Principles of Psychology by William James, Dover Publishing Company, 1890, pp. 593-616.

¹ It should be said that the methods of leaving the patient to himself, and that of the simple verbal suggestion of sleep (the so-called Nancy method introduced by Dr. Lié-

which they are told has been 'magnetized,' the drinking of 'magnetized' water, the reception of a letter ordering them to sleep, etc., are means which have been frequently employed. Recently M. Liégeois has hypnotized some of his subjects at a distance of $1\frac{1}{2}$ kilometres by giving them an intimation to that effect through a telephone. With some subjects, if you tell them in advance that at a certain hour of a certain day they will become entranced, the prophecy is fulfilled. Certain hysterical patients are immediately thrown into hypnotic catalepsy by any violent sensation, such as a blow on a gong or the flashing of an intense light in their eyes. Pressure on certain parts of the body (called *zones hypnogènes* by M. Pitres) rapidly produces hypnotic sleep in some hysterics. These regions, which differ in different subjects, are oftenest found on the forehead and about the root of the thumbs. Finally, persons in ordinary sleep may be transferred into the hypnotic condition by verbal intimation or contact, performed so gently as not to wake them up.

Some operators appear to be more

bault of that place), seem, wherever applicable, to be the best, as they entail none of the after-inconveniences which occasionally follow upon straining his eyes. A new patient should not be put through a great variety of different suggestions in immediate succession. He should be waked up from time to time, and then rehypnotized to avoid mental confusion and excitement. Before finally waking a subject you should *undo* whatever delusive suggestions you may have implanted in him, by telling him that they are all gone, etc., and that you are now going to restore him to his natural state. Headache, languor, etc., which sometimes follow the first trance or two, must be banished at the outset, by the operator strongly assuring the subject that such things *never* come from hypnotism, that the subject *must not* have them, etc.

successful than others in getting control of their subjects. I am informed that Mr. Gurney (who made valuable contributions to the theory of hypnotism) was never able himself to hypnotize, and had to use for his observations the subjects of others. On the other hand, Liébault claims that he hypnotizes 92 per cent of all comers, and Wetterstrand in Stockholm says that amongst 718 persons there proved to be only 18 whom he failed to influence. Some of this disparity is unquestionably due to differences in the personal 'authority' of the operator, for the prime condition of success is that the subject should confidently *expect* to be entranced. Much also depends on the operator's tact in interpreting the physiognomy of his subjects, so as to give the right commands, and 'crowd it on' to the subject, at just the propitious moments. These conditions account for the fact that operators grow more successful the more they operate. Bernheim says that whoever does not hypnotize 80 per cent of the persons whom he tries has not yet learned to operate as he should. Whether certain operators have over and above this a peculiar 'magnetic power' is a question which I leave at present undecided. Children under three or four, and insane persons, especially idiots, are unusually hard to hypnotize. This seems due to the impossibility of getting them to fix their attention continuously on the idea of the coming trance. All ages above infancy are probably equally hypnotizable, as are all races and both sexes. A certain amount of mental training, sufficient to aid concentration of the attention, seems a favorable condition, and so does a certain momentary indifference or passivity as to the result. Native strength or weakness of 'will' have absolutely

nothing to do with the matter. Frequent trances enormously increase the susceptibility of a subject, and many who resist at first succumb after several trials. Dr. Moll says he has more than once succeeded after forty fruitless attempts. Some experts are of the opinion that every one is hypnotizable essentially, the only difficulty being the more habitual presence in some individuals of hindering mental preoccupations, which, however, may suddenly at some moment be removed.

The trance may be dispelled instantaneously by saying in a rousing voice, 'All right, wake up!' or words of similar purport. At the Salpêtrière they awaken subjects by blowing on their eyelids. Upward passes have an awakening effect; sprinkling cold water ditto. Anything will awaken a patient who expects to be awakened by that thing. Tell him that he will wake after counting five, and he will do so. Tell him to wake in five minutes, and he is very likely to do so punctually, even though he interrupt thereby some exciting histrionic performance which you may have suggested.—As Dr. Moll says, any theory which pretends to explain the physiology of the hypnotic state must keep account of the fact that so simple a thing as hearing the word 'wake!' will end it.

The theory of Suggestion denies that there is any special hypnotic state worthy of the name of trance or neurosis. All the symptoms above described, as well as those to be described hereafter, are results of that mental susceptibility which we all to some degree possess, of yielding assent to outward suggestion, of affirming what we strongly conceive, and of acting in accordance with what we are

made to expect. The bodily symptoms of the Salpêtrière patients are all of them results of expectation and training. The first patients accidentally did certain things which their doctors thought typical and caused to be repeated. The subsequent subjects 'caught on' and followed the established tradition. In proof of this the fact is urged that the classical three stages and their grouped symptoms have *only* been reported as spontaneously occurring, so far, at the Salpêtrière, though they may be superinduced by deliberate suggestion, in patients anywhere found. The ocular symptoms, the flushed face, accelerated breathing, etc., are said not to be symptoms of the passage into the hypnotic state as such, but merely consequences of the strain on the eyes when the method of looking at a bright object is used. They are absent in the subjects at Nancy, where simple verbal suggestion is employed. The various reflex effects (aphasia, echolalia, imitation, etc.) are but habits induced by the influence of the operator, who unconsciously urges the subject into the direction in which he would prefer to have him go. The influence of the magnet, the opposite effects of upward and downward passes, etc., are similarly explained. Even that sleepy and inert condition, the advent of which seems to be the prime condition of farther symptoms being developed, is said to be merely due to the fact that the mind expects it to come; whilst its influence on the other symptoms is not physiological, so to speak, but psychological, its own easy realization by suggestion simply encouraging the subject to expect that ulterior suggestions will be realized with equal ease. The radical defenders of the suggestion-theory are thus led to deny the very existence of the

hypnotic state, in the sense of a peculiar trance-like condition which deprives the patient of spontaneity and makes him passive to suggestion from without. The trance itself is only one of the suggestions, and many subjects in fact can be made to exhibit the other hypnotic phenomena without the preliminary induction of this one.

The theory of suggestion may be said to be quite triumphant at the present day over the neurosis-theory as held at the Salpêtrière, with its three states, and its definite symptoms supposed to be produced by physical agents apart from co-operation of the subject's mind. But it is one thing to say this, and it is quite another thing to say that there is no peculiar physiological condition whatever worthy of the name of hypnotic trance, no peculiar state of nervous equilibrium, 'hypotaxy,' 'dissociation,' or whatever you please to call it, during which the subject's susceptibility to outward suggestion is greater than at ordinary times. All the facts seem to prove that, until this trance-like state is assumed by the patient, suggestion produces very insignificant results, but that, when it is once assumed, there are no limits to suggestion's power. The state in question has many affinities with ordinary sleep. It is probable, in fact, that we all pass through it transiently whenever we fall asleep; and one might most naturally describe the usual relation of operator and subject by saying that the former keeps the latter suspended between waking and sleeping by talking to him enough to keep his slumber from growing profound, and yet not in such a way as to wake him up. A hypnotized patient, *left to himself*, will either fall sound asleep or wake up entirely. The difficulty in hypnotizing refractory persons is that of catching them at the

right moment of transition and making it permanent. Fixing the eyes and relaxing the muscles of the body produce the hypnotic state just as they facilitate the advent of sleep. The first stages of ordinary sleep are characterized by a peculiar dispersed attitude of the attention. Images come before consciousness which are entirely incongruous with our ordinary beliefs and habits of thought. The latter either vanish altogether or withdraw, as it were, inertly into the background of the mind, and let the incongruous images reign alone. These images acquire, moreover, an exceptional vivacity; they become first 'hypnagogic hallucinations,' and then, as the sleep grows deeper, dreams. Now the 'monoidism,' or else the impotency and failure to 'rally' on the part of the background-ideas, which thus characterize somnolence, are unquestionably the result of a special physiological change occurring in the brain at that time. Just so that similar monoidism, or dissociation of the reigning fancy from those other thoughts which might possibly act as its 'reductives,' which characterize the hypnotic consciousness, must equally be due to a special cerebral change. The term 'hypnotic trance,' which I employ, tells us nothing of what the change is, but it marks the fact that it exists, and is consequently a useful expression. The great vivacity of the hypnotic images (as gauged by their motor effects), the oblivion of them when normal life is resumed, the abrupt awakening, the recollection of them again in subsequent trances, the anæsthesia and hyperæsthesia which are so frequent, all point away from our simple waking credulity and 'suggestibility' as the type by which the phenomena are to be interpreted, and make us look rather towards sleep and

dreaming, or towards those deeper alterations of the personality known as automatism, double consciousness, or 'second' personality for the true analogues of the hypnotic trance.² Even the best hypnotic subjects pass through life without any one suspecting them to possess such a remarkable susceptibility, until by deliberate experiment it is made manifest. The operator fixes their eyes or their attention a short time to develop the propitious phase, holds them in it by his talk, and *the state being there*, makes them the puppets of all his suggestions. But no ordinary suggestions of waking life ever took such control of their mind.

The suggestion-theory may therefore be approved as correct, provided we grant the trance-state as its prerequisite. The three states of Charcot, the strange reflexes of Heidenhain, and all the other bodily phenomena which have been called direct consequences of the trance-state itself, are not such. They are products of suggestion, the trance-state having no particular outward symptoms of its own; but without the trance-state there, those particular suggestions could never have been successfully made.³

² The state is not *identical* with sleep, however analogous in certain respects. The lighter stages of it, particularly, differ from sleep and dreaming, inasmuch as they are characterized almost exclusively by *muscular* inabilities and compulsions, which are not noted in ordinary somnolence, and the *mind*, which is confused in somnolence, may be quite clearly conscious, in the lighter state of trance, of all that is going on.

³ The word 'suggestion' has been bandied about too much as if it explained all mysteries: When the subject obeys it is by reason of the 'operator's suggestion'; when he proves refractory it is in consequence of an 'auto-suggestion' which he has made to himself, etc., etc. What explains everything ex-

The Symptoms of the Trance

This accounts for the altogether indefinite array of symptoms which have been gathered together as characteristic of the hypnotic state. The law of habit dominates hypnotic subjects even more than it does waking ones. Any sort of personal peculiarity, any trick accidentally fallen into in the first instance by some one subject, may, by attracting attention, become stereotyped, serve as a pattern for imitation, and figure as the type of a school. The first subject trains the operator, the operator trains the succeeding subjects, all of them in perfect good faith conspiring together to evolve a perfectly arbitrary result. With the extraordinary perspicacity and subtlety of perception which subjects often display for all that concerns the operator with whom they are *en rapport*, it is hard to keep them ignorant of anything which he expects. Thus it happens that one easily verifies on new subjects what one has already seen on old ones, or any desired symptom of which one may have heard or read.

The symptoms earliest observed by writers were all thought to be typical. But with the multiplication of observed phenomena, the importance of most particular symptoms as marks of the state has diminished. This lightens very much our own immediate task. Proceeding to enumerate the symptoms of the hypnotic trance, I

plains nothing; and it must be remembered that what *needs* explanation here is the fact that in a certain condition of the subject suggestions operate as they do *at no other time*; that through them functions are affected which ordinarily elude the action of the waking will; and that usually all this happens in a condition of which no after-memory remains.

may confine myself to those which are intrinsically interesting, or which differ considerably from the normal functions of man.

First of all comes *amnesia*. In the earlier stages of hypnotism the patient remembers what has happened, but with successive sittings he sinks into a deeper condition, which is commonly followed by complete loss of memory. He may have been led through the liveliest hallucinations and dramatic performances, and have exhibited the intensest apparent emotion, but on waking he can recall nothing at all. The same thing happens on waking from sleep in the midst of a dream—it quickly eludes recall. But just as we may be *reminded* of it, or of parts of it, by meeting persons or objects which figured therein, so on being adroitly prompted, the hypnotic patient will often remember what happened in his trance. One cause of the forgetfulness seems to be the disconnection of the trance performances with the system of waking ideas. Memory requires a continuous train of association. M. Delboëuf, reasoning in this way, woke his subjects in the midst of an action begun during trance (washing the hands, e.g.), and found that they then remembered the trance. The act in question bridged over the two states. But one can often make them remember by merely telling them during the trance that they *shall* remember. Acts of one trance, moreover, are usually recalled, either spontaneously or at command, during another trance, provided that the contents of the two trances be not mutually incompatible.

SUGGESTIBILITY

The patient believes everything which his hypnotizer tells him, and

does everything which the latter commands. Even results over which the will has normally no control, such as sneezing, secretion, reddening and growing pale, alterations of temperature and heart-beat, menstruation, action of the bowels, etc., may take place in consequence of the operator's firm assertions during the hypnotic trance, and the resulting conviction on the part of the subject, that the effects will occur. Since almost all the phenomena yet to be described are effects of this heightened suggestibility, I will say no more under the general head, but proceed to illustrate the peculiarity in detail.

Effects on the voluntary muscles seem to be those most easily got; and the ordinary routine of hypnotizing consists in provoking them first. Tell the patient that he cannot open his eyes or his mouth, cannot unclasp his hands or lower his raised arm, cannot rise from his seat, or pick up a certain object from the floor, and he will be immediately smitten with absolute impotence in these regards. The effect here is generally due to the *involuntary contraction* of antagonizing muscles. But one can equally well suggest *paralysis*, of an arm for example, in which case it will hang perfectly placid by the subject's side. Cataleptic and tetanic rigidity are easily produced by suggestion, aided by handling the parts. One of the favorite shows at public exhibitions is that of a subject stretched stiff as a board with his head on one chair and his heels on another. The cataleptic retention of impressed attitudes differs from voluntary assumption of the same attitude. An arm voluntarily held out straight will drop from fatigue after a quarter of an hour at the utmost, and before it falls the agent's distress will be made manifest

by oscillations in the arm, disturbances in the breathing, etc. But Charcot has shown that an arm held out in hypnotic catalepsy, though it may as soon descend, yet does so slowly and with no accompanying vibration, whilst the breathing remains entirely calm. He rightly points out that this shows a profound physiological change, and is proof positive against simulation, as far as this symptom is concerned. A cataleptic attitude, moreover, may be held for many hours.—Sometimes an expressive attitude, clinching of the fist, contraction of the brows, will gradually set up a sympathetic action of the other muscles of the body, so that at last a *tableau vivant* of fear, anger, disdain, prayer, or other emotional condition, is produced with rare perfection. This effect would seem to be due to the suggestion of the mental state by the first contraction. Stammering, aphasia, or inability to utter certain words, pronounce certain letters, are readily producible by suggestion.

Hallucinations of all the senses and *delusions* of every conceivable kind can be easily suggested to good subjects. The emotional effects are then often so lively, and the pantomime display so expressive, that it is hard not to believe in a certain 'psychic hyper-excitability,' as one of the concomitants of the hypnotic condition. You can make the subject think that he is freezing or burning, itching or covered with dirt, or wet; you can make him eat a potato for a peach, or drink a cup of vinegar for a glass of champagne;⁴ ammonia will smell to him like cologne water; a chair will

be a lion, a broom-stick a beautiful woman, a noise in the street will be an orchestral music, etc., etc., with no limit except your powers of invention and the patience of the lookers on.⁵ Illusions and hallucinations form the *pièces de résistance* at public exhibitions. The comic effect is at its climax when it is successfully suggested to the subject that his personality is changed into that of a baby, of a street boy, of a young lady dressing for a party, of a stump orator, or of Napoleon the Great. He may even be transformed into a beast, or an inanimate thing like a chair or a carpet, and in every case will act out all the details of the part with a sincerity and intensity seldom seen at the theatre. The excellence of the performance is in these cases the best reply to the suspicion that the subject may be shamming—so skilful a shammer must long since have found his true function in life upon the stage. Hallucinations and histrionic delusions generally go with a certain depth of the trance, and are followed by complete forgetfulness. The subject awakens from them at the command of the operator with a sudden start of surprise, and may seem for a while a little dazed.

Subjects in this condition will receive and execute suggestions of crime, and act out a theft, forgery,

⁴ A complete fit of drunkenness may be the consequence of the suggested champagne. It is even said that real drunkenness has been cured by suggestion.

⁵ The suggested hallucination may be followed by a negative after-image, just as if it were a real object. This can be very easily verified with the suggested hallucination of a colored cross on a sheet of white paper. The subject, on turning to another sheet of paper, will see a cross of the complementary color. Hallucinations have been shown by MM. Binet and Féré to be doubled by a prism or mirror, magnified by a lens, and in many other ways to behave optically like real objects. These points have been discussed already on p. 128 ff.

arson, or murder. A girl will believe that she is married to her hypnotizer, etc. It is unfair, however, to say that in these cases the subject is a pure puppet with no spontaneity. His spontaneity is certainly not in abeyance so far as things go which are harmoniously associated with the suggestion given him. He takes the text from his operator; but he may amplify and develop it enormously as he acts it out. His spontaneity is lost only for those systems of ideas which *conflict* with the suggested delusion. The latter is thus 'systematized'; the rest of consciousness is shut off, excluded, dissociated from it. In extreme cases the rest of the mind would seem to be actually abolished and the hypnotic subject to be literally a changed personality, a being in one of those 'second' states which we studied earlier. But the reign of the delusion is often not as absolute as this. If the thing suggested be too intimately repugnant, the subject may strenuously resist and get nervously excited in consequence, even to the point of having an hysterical attack. The conflicting ideas slumber in the background and merely permit those in the foreground to have their way until a *real* emergency arises; then they assert their rights. As M. Delbœuf says, the subject surrenders himself good-naturedly to the performance, stabs with the pasteboard dagger you give him because he knows what it is, and fires off the pistol because he knows it has no ball; but for a real murder he would not be your man. It is undoubtedly true that subjects are often well aware that they are acting a part. They know that what they do is absurd. They know that the hallucination which they see, describe, and act upon, is not really there. They may laugh at themselves; and they always

recognize the abnormality of their state when asked about it, and call it 'sleep.' One often notices a sort of mocking smile upon them, as if they were playing a comedy, and they may even say on 'coming to' that they were shamming all the while. These facts have misled ultra-skeptical people so far as to make them doubt the genuineness of any hypnotic phenomena at all. But, save the consciousness of 'sleep,' they do not occur in the deeper conditions; and when they do occur they are only a natural consequence of the fact that the 'mono-ideism' is incomplete. The background-thoughts still exist, and have the power of *comment* on the suggestions, but no power to inhibit their motor and associative effects. A similar condition is frequent enough in the waking state, when an impulse carries us away and our 'will' looks on wonderingly like an impotent spectator. These 'shammers' continue to sham in just the same way, every new time you hypnotize them, until at last they are forced to admit that if shamming there be, it is something very different from the free voluntary shamming of waking hours.

Real sensations may be abolished as well as false ones suggested. Legs and breasts may be amputated, children born, teeth extracted, in short the most painful experiences undergone, with no other anæsthetic than the hypnotizer's assurance that no pain shall be felt. Similarly morbid pains may be annihilated, neuralgias, toothaches, rheumatisms cured. The sensation of hunger has thus been abolished, so that a patient took no nourishment for fourteen days. The most interesting of these suggested anæsthesias are those limited to certain objects of perception. Thus a subject may be made blind to a certain person and to him alone, or deaf

to certain words but to no others.⁶ In this case the anæsthesia (or *negative hallucination*, as it has been called) is apt to become *systematized*. Other things related to the person to whom one has been made blind may also be shut out of consciousness. What he says is not heard, his contact is not felt, objects which he takes from his pocket are not seen, etc. Objects which he screens are seen as if he were transparent. Facts about him are forgotten, his name is not recognized when pronounced. Of course there is great variety in the completeness of this systematic extension of the suggested anæsthesia, but one may say that some tendency to it always exists. When one of the subjects' own limbs is made anæsthetic, for example, memories as well as sensations of its movements often seem to depart. An interesting degree of the phenomenon is found in the case related by M. Binet of a subject to whom it was suggested that a certain M. C. was invisible. She still saw M. C., but saw him as a stranger, having lost the memory of his name and his existence.—Nothing is easier than to make subjects forget their own name and condition in life. It is one of the suggestions which most promptly succeed, even with quite fresh ones. A systematized amnesia of certain periods of one's life may also be suggested, the subject placed, for instance, where he was a decade ago with the intervening years obliterated from his mind.

The mental condition which accompanies these systematized anæsthesias and amnesias is a very curious

one. The anæsthesia is not a genuine sensorial one, for if you make a real red cross (say) on a sheet of white paper invisible to an hypnotic subject, and yet cause him to look fixedly at a dot on the paper on or near the cross, he will, on transferring his eye to a blank sheet, see a bluish-green after-image of the cross. This proves that it has impressed his sensibility. He has *felt* it, but not *perceived* it. He had actively ignored it, refused to recognize it, as it were. Another experiment proves that he *must distinguish* it first in order thus to ignore it. Make a stroke on paper or black-board, and tell the subject it is not there, and he will see nothing but the clean paper or board. Next, he not looking, surround the original stroke with other strokes exactly like it, and ask him what he sees. He will point out one by one all the new strokes and omit the original one every time, no matter how numerous the new strokes may be, or in what order they are arranged. Similarly, if the original single stroke to which he is blind be *doubled* by a prism of sixteen degrees placed before one of his eyes (both being kept open), he will say that he now sees *one* stroke, and point in the direction in which the image seen through the prism lies.

Obviously, then, he is not blind to the *kind* of stroke in the least. He is blind only to one individual stroke of that kind in a particular position on the board or paper,—that is, to a particular complex object; and, paradoxical as it may seem to say so, he must distinguish it with great accuracy from others like it, in order to remain blind to it when the others are brought near. He 'apperceives' it, as a preliminary to not seeing it at all! How to conceive of this state of mind is not easy. It would be much simpler

⁶ M. Liégeois explains the common exhibition-trick of making the subject unable to get his arms into his coat-sleeves again after he has taken his coat off, by an anæsthesia to the necessary parts of the coat.

to understand the process, if adding new strokes made the first one visible. There would then be two different objects apperceived as totals,—paper with one stroke, paper with two strokes; and, blind to the former, he would see all that was *in* the latter, because he would have apperceived it as a different total in the first instance.

A process of this sort occurs sometimes (not always) when the new strokes, instead of being mere repetitions of the original one, are lines which combine with it into a total object, say a human face. The subject of the trance then may regain his sight of the line to which he had previously been blind, by seeing it as part of the face.

When by a prism before one eye a previously invisible line has been made visible to that eye, and the other eye is closed or screened, *its* closure makes no difference; the line still remains visible. But if *then* the prism is removed, the line will disappear even to the eye which a moment ago saw it, and both eyes will revert to their original blind state.

We have, then, to deal in these cases neither with a sensorial anæsthesia, nor with a mere failure to notice, but with something much more complex; namely, an active counting out and positive exclusion of certain objects. It is as when one 'cuts' an acquaintance, 'ignores' a claim, or 'refuses to be influenced' by a consideration of whose existence one remains aware. Thus a lover of Nature in America finds himself able to overlook and ignore entirely the board- and rail-fences and general roadside raggedness, and revel in the beauty and picturesqueness of the other elements of the landscape, whilst to a newly-arrived European the fences

are so aggressively present as to spoil enjoyment.

Messrs. Gurney, Janet, and Binet have shown that the ignored elements are preserved in a split-off portion of the subjects' consciousness which can be tapped in certain ways, and made to give an account of itself (see Vol. I. p. 209).

Hyperæsthesia of the senses is as common a symptom as anæsthesia. On the skin two points can be discriminated at less than the normal distance. The sense of touch is so delicate that (as M. Delbœuf informs me) a subject after simply poising on her finger-tips a blank card drawn from a pack of similar ones can pick it out from the pack again by its 'weight.' We approach here the line where, to many persons, it seems as if something more than the ordinary senses, however sharpened, were required in explanation. I have seen a coin from the operator's pocket repeatedly picked out by the subject from a heap of twenty others,⁷ by its greater 'weight' in the subject's language.—Auditory hyperæsthesia may enable a subject to hear a watch tick, or his operator speak, in a distant room.—One of the most extraordinary examples of visual hyperæsthesia is that reported by Bergson, in which a subject who seemed to be reading through the back of a book held and looked at by the operator, was really proved to be reading the image of the page reflected on the latter's cornea. The same subject was able to discriminate with the naked eye details in a microscopic preparation. Such cases of 'hyperæsthesia of vision' as that reported by Taguet and Sauvaire, where subjects could see things mirrored by non-

⁷ Precautions being taken against differences of temperature and other grounds of suggestion.

reflecting bodies, or through opaque pasteboard, would seem rather to belong to 'psychical research' than to the present category.—The ordinary test of visual hyperacuteness in hypnotism is the favorite trick of giving a subject the hallucination of a picture on a blank sheet of card-board, and then mixing the latter with a lot of other similar sheets. The subject will always find the picture on the original sheet again, and recognize infallibly if it has been turned over, or upside down, although the bystanders have to resort to artifice to identify it again. The subject notes peculiarities on the card, too small for waking observation to detect.⁸ If it be said that the spectators guide him by their manner, their breathing, etc., that is only another proof of his hyperæsthesia; for he undoubtedly is conscious of subtler personal indications (of his operator's mental states especially) than he could notice in his waking state. Examples of this are found in the so-called 'magnetic rapport.' This is a name for the fact that in deep trance, or in lighter trance whenever the suggestion is made, the subject is deaf and blind to everyone but the operator or those spectators to whom the latter expressly awakens his senses. The most violent appeals from anyone else are for him as if non-existent, whilst he obeys the faintest signals on the part of his hypnotizer. If in catalepsy, his limbs will retain their attitude only when the operator moves them; when others move them they fall down, etc. A more remarkable fact still is that the patient will often answer anyone

whom his operator touches, or at whom he even points his finger, in however concealed a manner. All which is rationally explicable by expectation and suggestion, if only it be farther admitted that his senses are acutely sharpened for all the operator's movements.⁹ He often shows great anxiety and restlessness if the latter is out of the room. A favorite experiment of Mr. E. Gurney's was to put the subject's hands through an opaque screen, and cause the operator to point at one finger. *That* finger presently grew insensible or rigid. A bystander pointing simultaneously at another finger, never made that insensible or rigid. Of course the elective *rappport* with their operator had been developed in these trained subjects during the hypnotic state, but the phenomenon then occurred in some of them during the waking state, even when their consciousness was absorbed in animated conversation with a fourth party.¹⁰ I confess that when I saw these experiments I was impressed with the necessity for admitting between the *emanations* from different people differences for which we have no name, and a discriminative sensibility for them of the nature of which we can form no clear conception, but which seems to be developed in certain subjects by the hypnotic trance.—The enigmatic reports of

⁹ I must repeat, however, that we are here on the verge of possibly unknown forces and modes of communication. Hypnotization at a distance, with no grounds for expectation on the subject's part that it was to be tried, seems pretty well established in certain very rare cases. See in general, for information on these matters, the *Proceedings of the Soc. for Psych. Research, passim*.

¹⁰ Here again the perception in question must take place below the threshold of ordinary consciousness, possibly in one of those split-off selves or 'second' states whose existence we have so often to recognize.

⁸ It should be said, however, that the bystander's ability to discriminate unmarked cards and sheets of paper from each other is much greater than one would naturally suppose.

the effect of magnets and metals, even if they be due, as many contend, to unintentional suggestion on the operator's part, certainly involve hyperæsthetic perception, for the operator seeks as well as possible to conceal the moment when the magnet is brought into play, and yet the subject not only finds it out that moment in a way difficult to understand, but may develop effects which (in the first instance certainly) the operator did not expect to find. Unilateral contractions, movements, paralyses, hallucinations, etc., are made to pass to the other side of the body, hallucinations to disappear, or to change to the complementary color, suggested emotions to pass into their opposites, etc. Many Italian observations agree with the French ones, and the upshot is that if unconscious suggestion lie at the bottom of this matter, the patients show an enormously exalted power of divining what it is they are expected to do. This hyperæsthetic perception is what concerns us now.¹¹ Its *modus* cannot yet be said to be defined.

Changes in the nutrition of the tissues may be produced by suggestion. These effects lead into therapeutics—a subject which I do not propose to treat of here. But I may say that there seems no reasonable ground for doubting that in certain chosen sub-

jects the suggestion of a congestion, a burn, a blister, a raised papule, or a bleeding from the nose or skin, may produce the effect. Messrs. Beaunis, Berjon, Bernheim, Bourru, Burot, Charcot, Delbœuf, Dumontpallier, Focachon, Forel, Jendrassik, Krafft-Ebing, Liébault, Liégeois, Lipp, Mabile, and others have recently vouched for one or other of these effects. Messrs. Delbœuf and Liégeois have annulled by suggestion, one the effects of a burn, the other of a blister. Delbœuf was led to his experiments after seeing a burn on the skin produced by suggestion, at the Salpêtrière, by reasoning that if the idea of a pain could produce inflammation it must be because pain was itself an inflammatory irritant, and that the abolition of it from a real burn ought therefore to entail the absence of inflammation. He applied the actual cautery (as well as vesicants) to symmetrical places on the skin, affirming that no pain should be felt on one of the sides. The result was a dry scorch on that side, with (as he assures me) no aftermark, but on the other side a regular blister with suppuration and a subsequent scar. This explains the innocuity of certain assaults made on subjects during trance. To test simulation, recourse is often had to sticking pins under their finger-nails or through their tongue, to inhalations of strong ammonia, and the like. These irritations, when not felt by the subject, seem to leave no after-consequences. One is reminded of the reported non-inflammatory character of the wounds made on themselves by dervishes in their pious orgies. On the other hand, the reddenings and bleedings of the skin along certain lines, suggested by tracing lines or pressing objects thereupon, put the accounts handed down to us of the

¹¹ I myself verified many of the above effects of the magnet on a blindfolded subject on whom I was trying them for the first time, and whom I believe to have never heard of them before. The moment, however, an opaque screen was added to the blindfolding, the effects ceased to coincide with the approximation of the magnet, so that it looks as if visual perception had been instrumental in producing them. The subject passed from my observation, so that I never could clear up the mystery. Of course I gave him consciously no hint of what I was looking for.

stigmata of the cross appearing on the hands, feet, sides, and forehead of certain Catholic mystics in a new light. As so often happens, a fact is denied until a welcome interpretation comes with it. Then it is admitted readily enough; and evidence judged quite insufficient to back a claim, so long as the church had an interest in making it, proves to be quite sufficient for modern scientific enlightenment, the moment it appears that a reputed saint can thereby be classed as 'a case of hystero-epilepsy.'

There remain two other topics, viz., post-hypnotic effects of suggestion, and effects of suggestion in the waking state.

Post-hypnotic, or deferred, suggestions are such as are given to the patients during trance, to take effect after waking. They succeed with a certain number of patients even when the execution is named for a remote period—months or even a year, in one case reported by M. Liégeois. In this way one can make the patient feel a pain, or be paralyzed, or be hungry or thirsty, or have an hallucination, positive or negative, or perform some fantastic action after emerging from his trance. The effect in question may be ordered to take place not immediately, but after an interval of time has elapsed, and the interval may be left to the subject to measure, or may be marked by a certain signal. The moment the signal occurs, or the time is run out, the subject, who until then seems in a perfectly normal waking condition, will experience the suggested effect. In many instances, whilst thus obedient to the suggestion, he seems to fall into the hypnotic condition again. This is proved by the fact that the moment the hallucination or suggested performance is over

he forgets it, denies all knowledge of it, and so forth; and by the further fact that he is 'suggestible' during its performance, that is, will receive new hallucinations, etc., at command. A moment later and this suggestibility has disappeared. It cannot be said, however, that relapse into the trance is an absolutely necessary condition for the post-hypnotic carrying out of commands, for the subject may be neither suggestible nor amnesic, and may struggle with all the strength of his will against the absurdity of this impulse which he feels rising in him, he knows not why. In these cases, as in most cases, he forgets the circumstance of the impulse having been suggested to him in a previous trance; regards it as arising within himself; and often improvises, as he yields to it, some more or less plausible or ingenious motive by which to justify it to the lookers-on. He acts, in short, with his usual sense of personal spontaneity and freedom; and the disbelievers in the freedom of the will have naturally made much of these cases in their attempts to show it to be an illusion.

The only really mysterious feature of these deferred suggestions is the patient's absolute ignorance during the interval preceding their execution that they have been deposited in his mind. They will often surge up at the preappointed time, even though you have vainly tried a while before to make him recall the circumstances of their production. The most important class of post-hypnotic suggestions are, of course, those relative to the patient's health—bowels, sleep, and other bodily functions. Among the most *interesting* (apart from the hallucinations) are those relative to future trances. One can determine the

hour and minute, or the signal, at which the patient will of his own accord lapse into trance again. One can make him susceptible in future to another operator who may have been unsuccessful with him in the past. Or more important still in certain cases, one can, by suggesting that certain persons shall never be able hereafter to put him to sleep, remove him for all future time from hypnotic influences which might be dangerous. This, indeed, is the simple and natural safeguard against those 'dangers of hypnotism' of which uninstructed persons talk so vaguely. A subject who knows himself to be ultra-susceptible should never allow himself to be entranced by an operator in whose moral delicacy he lacks complete confidence; and he can use a trusted operator's suggestions to protect himself against liberties which others, knowing his weakness, might be tempted to take with him.

The mechanism by which the command is retained until the moment for its execution arrives is a mystery which has given rise to much discussion. The experiments of Gurney and the observations of M. Pierre Janet and others on certain hysterical somnambulists seem to prove that it is stored up in consciousness; not simply organically registered, but that *the consciousness which thus retains it is split off, dissociated from the rest of the subject's mind*. We have here, in short, an experimental production of one of those 'second' states of the personality of which we have spoken so often. Only here the second state coexists as well as alternates with the first. Gurney had the brilliant idea of *tapping* this second consciousness by means of the planchette. He found that certain persons, who were both

hypnotic subjects and automatic writers, would if their hands were placed on a planchette (after being wakened from a trance in which they had received the suggestion of something to be done at a later time) write out unconsciously the order, or something connected with it. This shows that something inside of them, which could express itself through the hand alone, was continuing to think of the order, and possibly of it alone. These researches have opened a new vista of possible experimental investigations into the so-called 'second' states of the personality.

Some subjects seem almost as obedient to suggestion in the waking state as in sleep, or even more so, according to certain observers. Not only muscular phenomena, but changes of personality and hallucinations are recorded as the result of simple affirmation on the operator's part, without the previous ceremony of 'magnetizing' or putting into the 'mesmeric sleep.' These are all trained subjects, however, so far as I know, and the affirmation must apparently be accompanied by the patient concentrating his attention and gazing, however briefly, into the eyes of the operator. It is probable therefore that an extremely rapidly induced condition of trance is a prerequisite for success in these experiments.

I have now made mention of all the more important phenomena of the hypnotic trance. Of their therapeutic or forensic bearings this it not the proper place to speak. The recent literature of the subject is quite voluminous, but much of it consists in repetition. The best compendious work on the subject is 'Der Hypnotismus,' by Dr. A. Moll (Berlin, 1889; and just translated into English,

N. Y., 1890), which is extraordinarily complete and judicious. The other writings most recommendable are subjoined in the note.¹² Most of them

contain a historical sketch and much bibliography. A complete bibliography has been published by M. Dessoir (Berlin, 1888).

¹² Binet and Féré, 'Animal Magnetism,' in the International Scientific Series; A. Bernheim, 'Suggestive Therapeutics' (N. Y., 1889); J. Liégeois, 'De la Suggestion' (1889); E. Gurney, two articles in *Mind*, vol. ix.—In the recent revival of interest in the history of this subject, it seems a pity that the admirably critical and scientific work of Dr. John

Kearsley Mitchell of Philadelphia should remain relatively so unknown. It is quite worthy to rank with Braid's investigations. See "Five Essays" by the above author, edited by S. Weir Mitchell, Philadelphia, 1859, pp. 141-274.

Postscript to the chapter by James: This chapter is just as relevant today as it was in 1890, when it was first published. If you enjoy James's polished prose you should look in greater detail at the entire 1400-page work.

INTRODUCTION

What is the best way to suppress aggressive behavior? If you suggest counteraggression you're probably wrong. Azrin and his colleagues show that painful situations for an animal produce automatic aggressivity and violent behavior. This fascinating series of experiments raises many and socially relevant questions. For example, does war beget war? Read and decide. Does your own experience point in the same direction?

PAIN-INDUCED FIGHTING IN THE SQUIRREL MONKEY

N. H. Azrin / R. R. Hutchinson / D. F. Hake

Reflexive fighting in response to pain-shock has been obtained for mice (Tedeschi, 1959), hamsters (Ulrich and Azrin, 1962), cats (Ulrich, personal communication), and several strains of rats (Ulrich and Azrin, 1962). This experiment explored the possible existence of the reflex in primates.

Eight squirrel monkeys were subjects. An experimental space 36 by 24 by 18 in. high provided ample

room for the monkeys to move about without contacting each other. The floor of the chamber was constructed of parallel grids through which shocks could be delivered. A large one-way window allowed unrestricted observation. One pair of monkeys was placed in the chamber and observed for approximately 5 min. In the absence of shock, the two monkeys explored the cage, briefly sat in the corner and moved past each other showing no signs of fighting. Brief (.15 sec) shocks of 3 ma intensity were then delivered every 30 sec. Upon the first shock delivery, the monkeys jumped and squealed but made no movements at

Reprinted with permission of the authors and publisher from the Journal of Experimental Analysis of Behavior, 1963, 6, p. 620. Copyright © 1963 by the Society for the Experimental Analysis of Behavior, Inc.

mouth, thereby providing an objective record of the attack. The critical feature of this method was that an additional response was required: the ball had to be pulled as a prerequisite for attack to occur. A still better method would be to utilize a response requirement that was not a component part of the attack response, the nature of the response being dictated by methodological considerations rather than by the nature of the attack behavior. This objective of utilizing an arbitrary response to measure a motivational state is, of course, the same objective that has led to the use of bar-presses, chain-pulls, and panel-presses to measure the strength of motivational states as diverse as hunger and thirst (Skinner, 1938), imprinting (Peterson, 1960), escape (Dinsmoor and Hughes, 1956), avoidance (Sidman, 1953), and intracranial reinforcement (Olds and Milner, 1954). Among other advantages, the use of an arbitrary response also makes it possible to apply a common measure to various motivational states that otherwise are manifested by quite dissimilar behavior patterns. The present experiment attempts to ascertain whether an arbitrary response can be acquired by means of operant reinforcement when the opportunity to attack is utilized as the reinforcing event for that response.

Method

The general procedure in the present study was to use tail-shock to induce aggression in monkeys. An inanimate object was used as the object-of-attack.

SUBJECTS

Five experimentally naive male squirrel monkeys served. Three other

monkeys were discontinued: one because of an unusual adaptation to the shock which resulted in large intrasession changes in the probability of attack, and two because of the low probability of attack ($<.10$) against the inanimate object. This failure to elicit attack consistently from an occasional monkey has been noted previously (Azrin, Hutchinson, and Hake, 1963). Since the monkeys were not raised in captivity, many of their characteristics that are not completely known may have been contributing factors. For example, Hutchinson, Azrin, and Ulrich (in press) found that with rats, age, social isolation, and hormonal development were contributing factors for the existence of shock-elicited attack. The weights of the five subjects in the present study ranged from 570 to 870 g, with an average weight of 740 g. During the experiments, subjects were individually housed in cages in which food and water were continuously available.

APPARATUS

The experimental chamber (see Fig. 1) measured 10 by 6 by 23 in. high and included a special chair (Hake and Azrin, 1963) which held the monkey in a loosely restrained position while allowing the delivery of pain-shock through tail-electrodes. The chamber was enclosed in a sound-attenuating enclosure that contained a one-way mirror. The tail-shock was delivered at an intensity of 400 v ac from the secondary of a transformer and through a 10 K ohm series resistor that stabilized the current flow. Each shock was 100 msec in duration.

The top of the inner chamber had an opening through which a canvas-covered ball, 2 in. in diameter, could

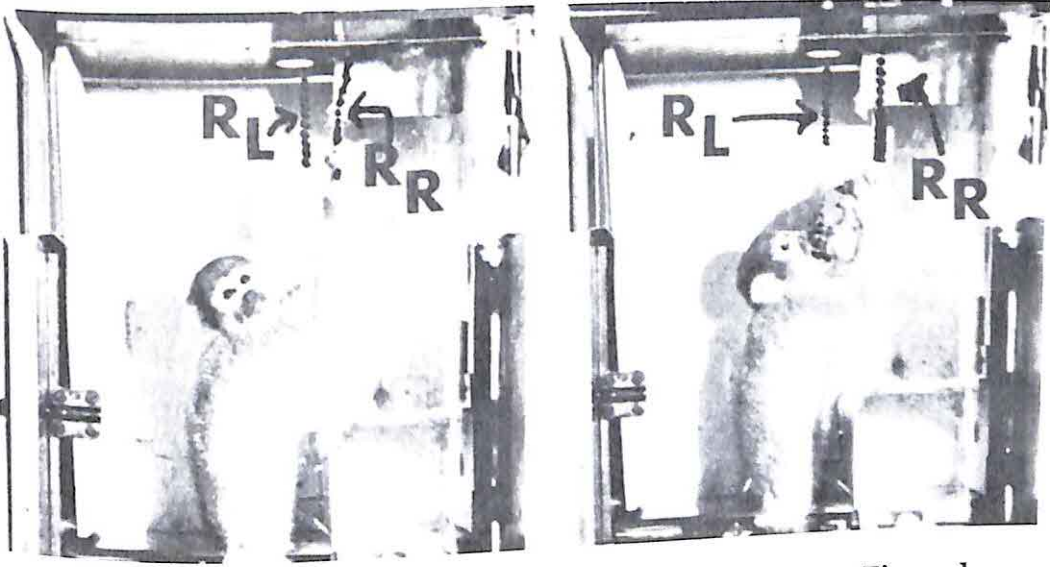


Figure 1

Left photograph: Squirrel monkey pulling on the R_R manipulandum. Right photograph: Attack on ball which has just been lowered as a result of a response on the R_R manipulandum.

be lowered by activating a motor. The ball could be similarly recovered through the opening by reversing the motor. "Reinforcement" as used here refers to the lowering of the ball through the opening to a distance of 4 in. from the ceiling where it remained for a duration of about 2 sec. At the end of the 2-sec period, the ball was automatically withdrawn through the opening. While the ball was in the lowered position, the monkey could grasp the ball and bring it to its mouth because of the flexible cord by which the ball was suspended. The cord (by which the ball was suspended) was attached to a micro-switch, thereby providing closure of the switch contacts when the monkey pulled the ball to its mouth. This switch closure required a minimum of 80 g of force.

The manipulanda for the conditioned response were two chains suspended through separate openings in the ceiling from two individual micro-

switches. The manipulandum on the right of the monkey is designated as R_R and that on the left as R_L . Both chains extended a distance of 4 in. from the ceiling, sufficient to enable the monkey to grasp the chain easily, yet not long enough to be disturbed by the subjects' casual movements. The two chains were 4.5 in. apart and were deliberately made physically different to increase the likelihood that subjects would discriminate between them. One of the manipulanda consisted of a simple bead chain; the other consisted of a bead chain enclosed in a narrow 0.25-in. diameter metal tube. Bead chain was used rather than the usual projecting bar or lever to reduce the subject's tendency to attack the manipulandum upon the delivery of the pain-shock. A pull exceeding 20 g was necessary to activate either switch; a response was defined as closure of the switch for a duration of about 0.5 sec. Maintained closure of this switch in excess

of 0.5 sec counted as only one response. To be counted as a second response, the chain had to be released and pulled a second time. The 0.5-sec requirement was imposed to eliminate the possibility of counting as responses any momentary closures of the switch that resulted from random movements of the monkey. An important aspect of the apparatus, as mentioned earlier in Azrin *et al.* (1964), is that there be no projecting objects in the chamber that could be attacked easily other than the intended object-of-attack, which was the ball. All of the walls, as well as the waistlock which restrained the monkey, were constructed of micarta, the hard smooth surface of which effectively discouraged any biting attacks. The left section of Fig. 1 is a photograph of a monkey responding to R_R . The right section of Fig. 1 is a photograph of the subject attacking the lowered ball after having just produced the ball by pulling the chain.

PROCEDURE

Table 1 outlines the experimental procedure. In Phase I, subjects were seated in the experimental chamber for 30 min. The tail was restrained in the electrode assembly, but no

shock was delivered. The ball was in a lowered position for the entire 30-min session. The purpose of this phase was to ascertain the frequency of attack against the ball when no shock was delivered. Biting and grabbing of the ball by the monkey was recorded (1) automatically by means of the microswitch that was connected to the ball and (2) manually by an observer through direct observation.

Phase II was identical to Phase I except that the brief tail-shocks were delivered every 15 sec for a total of 25 shocks. The purpose of Phase II was to ascertain whether attack would result against the ball upon the delivery of tail-shock when the ball was continuously present. This procedure was identical to that of a previous study (Azrin *et al.*, 1964).

During Phase III, only R_R was available. The brief shocks were delivered at regular intervals of 15 sec for a total of 120 shocks during the 30-min session. The ball could be lowered only by pulling the chain. Phase III attempted to ascertain whether the response of pulling R_R would be learned if that response were followed by the reinforcement of having the ball lowered for 2 sec.

In Phase IV, manipulandum R_L

TABLE 1

PHASE	SHOCK CONDITION	REINFORCEMENT SCHEDULE (AVAILABILITY OF BALL)	MANIPULANDUM AVAILABILITY
I	no shock	Continuously present	None
II	shock every 15 secs	Continuously present	None
III	shock every 15 secs	$R_R \rightarrow$ reinforcement	R_R
IV	shock every 15 secs	$R_L \rightarrow$ reinforcement	R_L
V	shock every 15 secs	$R_L \rightarrow$ reinforcement	R_L, R_R
VI	shock every 15 secs	$R_R \rightarrow$ reinforcement	R_L, R_R
VII	shock every 15 secs	$R_L \rightarrow$ reinforcement	R_L, R_R
VIIIa	shock every 15 secs	$R_R \rightarrow$ reinforcement	R_L, R_R
IX	no shock	$R_R \rightarrow$ reinforcement	R_L, R_R

aThis Phase omitted for one subject.

was available and R_R was absent. A closure of the R_L switch for 0.5 sec resulted in the immediate lowering of the ball for 2 sec similar to the procedures used in reinforcing R_R . Phase IV was designed to determine whether the monkey could also be conditioned to R_L . Phases III and IV provided reinforcement when only one manipulandum was present.

Phase V provided reinforcement for R_L when both R_R and R_L were available. This procedure is comparable to the common food reinforcement procedure in which two manipulanda are present, but responses on only one manipulandum result in the food reinforcement. Reinforcement on R_L was continued during successive sessions until the reinforced responses, R_L , constituted 80% or more of the total number of responses, $R_L + R_R$. In Phase VI, the reinforcement was discontinued for R_L and provided for R_R , thereby reversing the reinforcement contingencies. The number of sessions provided on Phase VI was determined by the same criterion used in Phase V. Upon reaching this criterion, the reinforcement contingency was then reversed in Phase VII, such that R_L was being reinforced and the responses on R_R were again nonreinforced just as they were in Phase V. During Phase VIII the reinforcement contingencies were again reversed.

The last Phase of the experiment, Phase IX, was identical to Phase VIII except that no shock was delivered.

The sessions were scheduled at the same time daily, Sundays excluded.

Results

Little or no attack occurred toward the continuously present ball in Phase I when no shocks were delivered.

Three subjects did bite or manipulate the ball during the first few minutes of the 30-min session, but by the end of 5 min, visual observations revealed no further contact.

When the shocks were delivered to the monkeys in Phase II while the ball was in the lowered position, attack occurred consistently and immediately after each delivery of the pain-shock. Only the first few shocks failed to elicit attack against the ball since the initial reaction of the monkey to the shock was typically a scrambling escape-like behavior frequently accompanied by unsuccessful attempts to bite or attack the walls of the chamber. By the 10th shock, however, each subject vigorously bit the lowered ball immediately upon the delivery of every pain-shock. This consistent elicitation of attack against an inanimate object by pain-shock is in agreement with previous results utilizing the same procedure (Azrin *et al.*, 1964).

When the availability of the ball was made contingent on the chain-pull (Phases III and IV), the chain-pulling responses occurred consistently after each shock. Figure 2 shows the temporal distribution of the R_L responses during Phase IV when an R_L response resulted in the delivery of the reinforcement (ball lowered). It can be seen that the R_L responses were emitted primarily within the first 3.0 sec after the shock delivery. All subjects averaged at least one chain-pulling response for each of the 120 shocks. The results (not shown) for Phase III showed a similar temporal distribution of the R_R responses.

When both manipulanda were present (Phases V-VIII) the monkey responded to that chain which produced the ball. Figure 3 shows the typical day-by-day results when

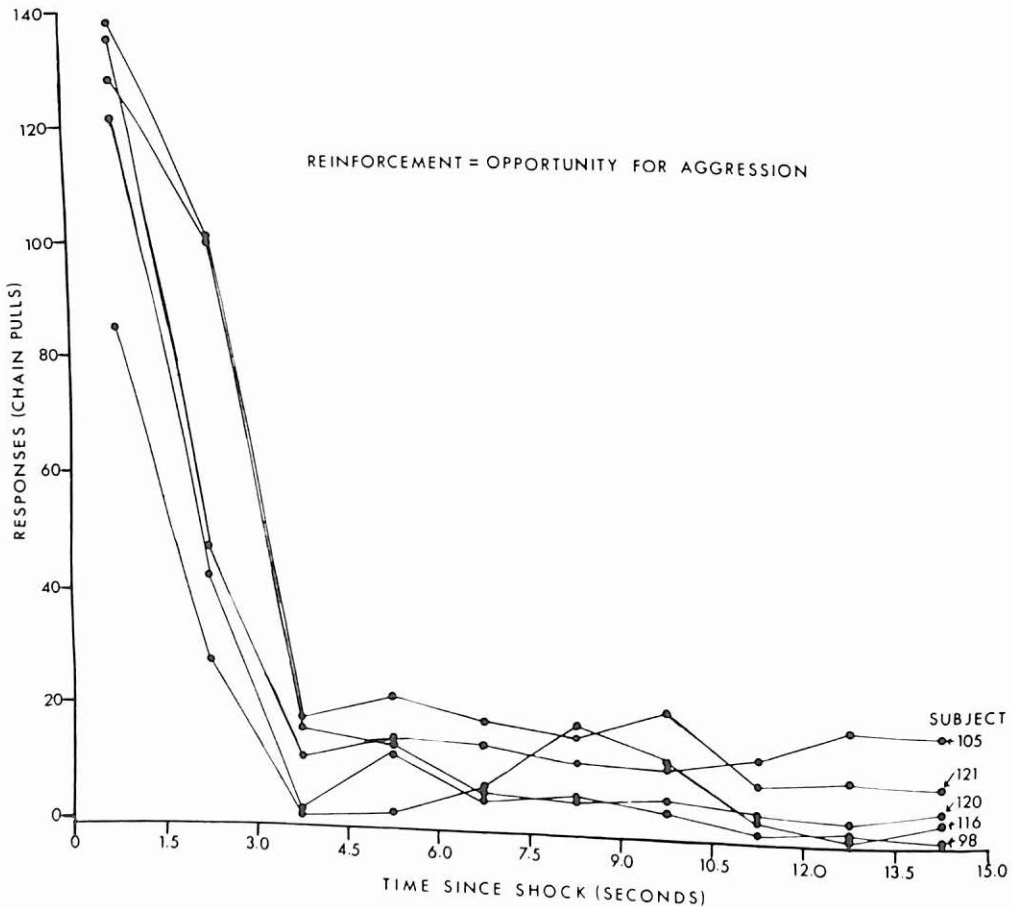


Figure 2
The number of chain-pulling responses as a function of time since the delivery of shock. The 15-sec period between shock deliveries was divided into 10 intervals of 1.5 sec. Each data point is the number of chain-pulling responses that occurred during each 1.5-sec class interval during one session.

the reinforcement was successively changed from R_L to R_R back to R_L then back to R_R . The frequency of each response increased when reinforcement was provided for that response and decreased when reinforcement was provided for the other response. The acquisition of the reinforced response and extinction of the nonreinforced response usually occurred within two or three sessions

for the other four subjects as well as for the subject depicted in Fig. 3.

All subjects showed this increase in frequency of the particular response that resulted in the availability of the object-of-attack, and a decrease in the frequency of the response that did not result in the availability of the object-of-attack. This is shown in Fig. 4, which compares the nonreinforced with the reinforced chain-pulling re-

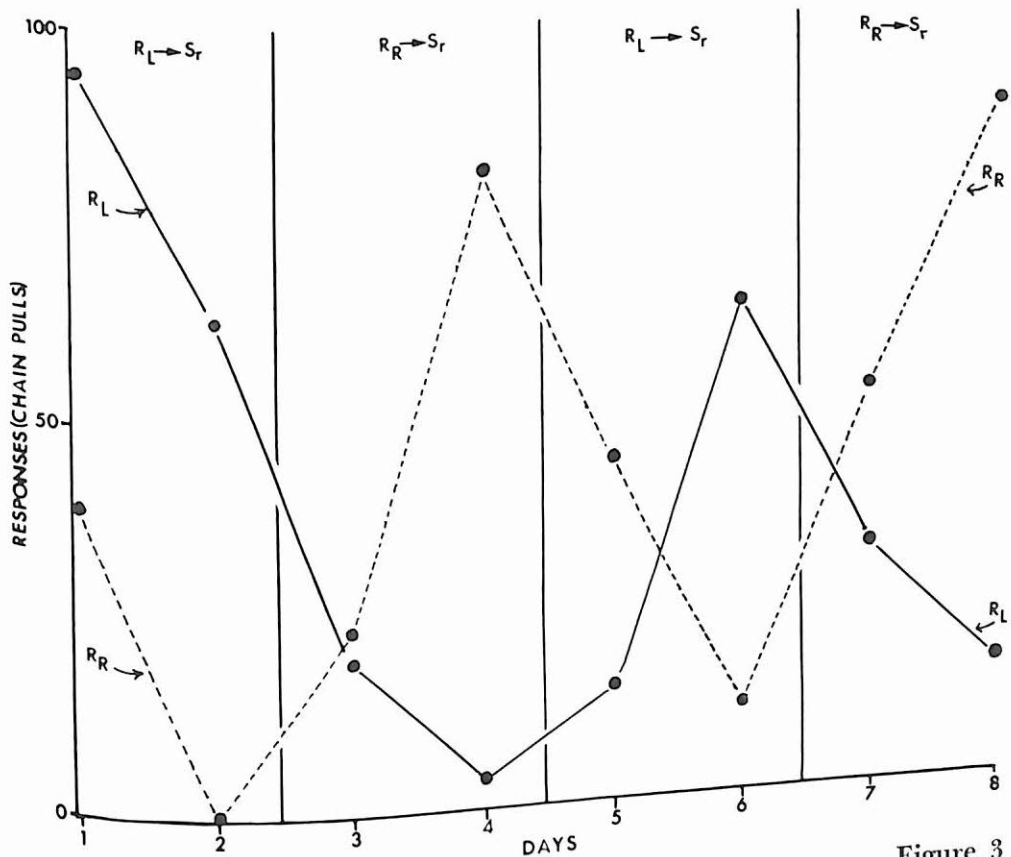
$S_r = \text{OPPORTUNITY FOR AGGRESSION}$ 

Figure 3

Reversal of the reinforcement contingency for one subject. The dotted line designates responses on the left chain (R_L); the solid line designates responses on the right chain (R_R). Each data point represents the number of responses per day. The reinforcement contingency was reversed at the point designated by the narrow vertical lines.

sponses for each of the five subjects. The data is presented in terms of the percentage of chain-pulling responses and is based on the final day of reinforcement for Phases V, VI, VII, and VIII that preceded the reversal of the reinforcement contingencies. Each of the scores is based on four sessions (three sessions for one subject). It can be seen that for all subjects, over 85%

of the chain-pulling responses were emitted on the chain that resulted in the delivery of the object-of-attack. The absolute number of responses emitted is indicated for each of the subjects.

Figure 5 compares the number of R_L responses emitted in the presence of shock with the number emitted in the absence of shock. The white bar

for each subject designates the number of R_L responses during the last session of Phase VIII, when the shocks were being presented every 15 sec. The solid bar designates the number of R_L responses during the single session of Phase IX when shock was not delivered. The R_L responses produced the ball under both conditions. It can be seen that the number of responses in the absence of shock was zero for subjects S98 and S105, and near-zero for the other three. Of the R_L responses that did occur in the absence of shock, all were emitted during the first 5 min of the 30-min session.

In the present procedure, reinforcement for a chain-pulling response was arranged by making an object-of-at-

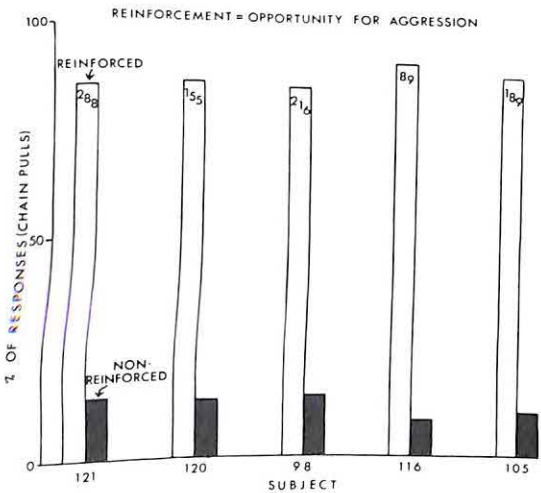


Figure 4

The average number and percentage of reinforced and nonreinforced chain-pulling responses for each subject. The R_L and R_R responses were combined. The absolute number of reinforced responses is presented at the top portion of the white bar.

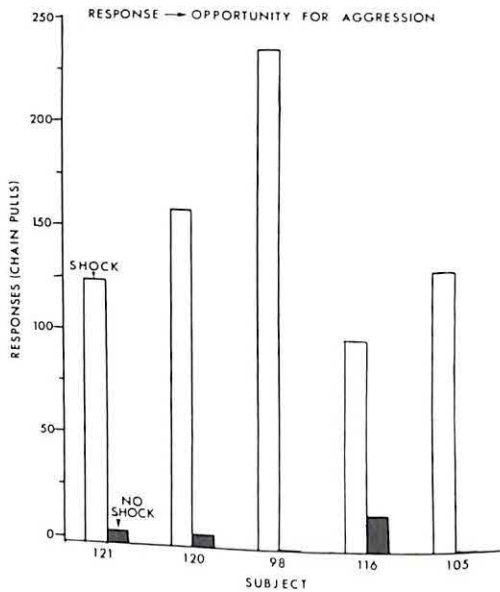


Figure 5

Chain-pulling responses for each subject in the presence and absence of shock. Each bar designates the responses for one day.

tack available for 2 sec. The question may be asked as to whether attack did indeed occur upon each presentation of the object-of-attack. An answer was attempted by taking two different measures of attack behavior. The first of these measures has already been described: a microswitch was attached to the cord by which the ball was suspended, so that the switch closure automatically registered on the recording equipment whenever the monkey pulled the object-of-attack toward its mouth. A second measure was obtained by visually observing each of the monkeys for 300 presentations of the object-of-attack during the course of the experiment. Each of these presentations resulted from a chain-pull. Figure 6 presents for each of the subjects the number of attacks as meas-

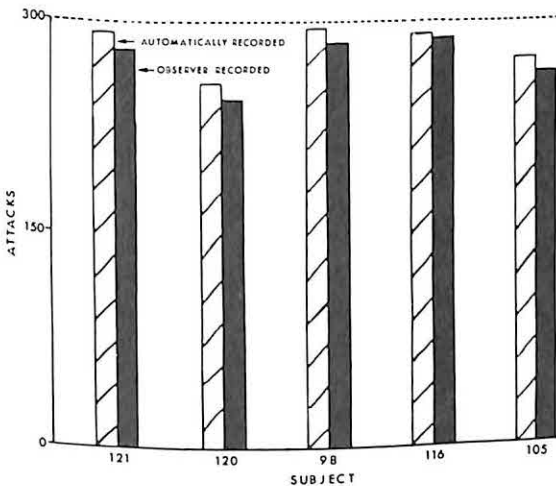


Figure 6

Comparison of two measures of attack based on 300 opportunities for attack.

ured by these two different methods of recording. The horizontal dotted line shows the total number of presentations (300) of the object-of-attack. The stippled bar shows the number of attacks as measured by the closure of the microswitch to which the object-of-attack was attached. No more than one attack was considered to have occurred during a given presentation of the object-of-attack even though the switch may have repeatedly closed and opened during that presentation. Thus, the number of attacks as measured by the closure of the switch could not exceed the number of presentations of the object-of-attack. It can be seen that the number of attacks as measured by the closure of the switch corresponds very closely to the number of presentations of the object-of-attack. For three of these subjects, the correspondence was almost perfect: a closure of the switch resulted on 97% or more of the presentations of the object-of-

attack. The correspondence for the other two was fairly close: 85% and 91%. The correspondence between the observer-noted attack (solid bar) and the automatic switch-recorded attack (stippled bar) is again very close, differing by no more than 5% for any subject. The observer-recorded attacks also correspond closely with the number of opportunities for attack: the greatest discrepancy is for S120, about 15%, but less than 10% for the other four subjects.

Discussion

The present findings confirmed the results of previous studies (Azrin *et al.*, 1964) in demonstrating the existence of attack against inanimate objects as a result of aversive stimulation. Little or no attack was seen in the present study when the object-of-attack was continuously available but no shock was presented. On the other hand, attack was consistently elicited toward the continuously available object-of-attack when shocks were presented. Similarly, the chain-pulling responses dropped to zero level in the absence of shock, but increased to a high level when the shocks were being delivered. Thus, it appears that the delivery of shock was necessary to produce aggression as measured either by the actual biting of the object-of-attack or by the frequency of the chain-pulling responses that produced the object-of-attack. The successful conditioning of the chain-pulling response offers a way to quantify the strength of aggression-motivation without the need to adapt the recording apparatus to the perhaps unique mode of attack (biting, grabbing, slapping, "threatening") that happens to be utilized by a par-

ticular animal or a particular species at a given moment.

Stimulus change appeared to contribute little to the occurrence of the chain-pulling response: a near zero level of chain-pulling was obtained in Phase IX when the response produced the ball but no shocks were delivered.

The chain-pulling response did not appear to be the result of a blind attack by the monkey upon the chain following the delivery of a shock since responses occurred almost exclusively only on that chain, which was followed by an operant consequence. It seems, therefore, that the principal source of strength of the chain-pulling response was the operant consequence of allowing an opportunity to attack the inanimate object.

It seems useful to consider aggression as a distinctive motivational state in the same loose sense in which it is useful to consider hunger or food deprivation as a distinctive motivational state. The opportunity to attack appears to be a reinforcement for a subject exposed to aversive stimulation in the same general sense that the opportunity to eat is a reinforcement for a subject that has been deprived of food. A new response was acquired and extinguished by arranging the opportunity to attack as a consequence for that response; similarly, new responses have been acquired and eliminated or extinguished by arranging the opportunity for eating. An obvious difference between aggression motivation and hunger motivation appears to be the relative brevity of the aggressive motivation. The aggressive tendencies reached a high level during the period immediately after the shock, falling off rapidly thereafter in contrast with hunger motivation, which persists for long periods of time. However, this

difference between aggression and hunger motivation may be only apparent and may derive primarily from the technical difficulty of maintaining a continuing state of aversive stimulation in a manner fairly comparable to the maintenance of a continuing state of food deprivation. Some support for this suggestion was obtained in previous studies. When continuous foot-shock was provided for short durations (Azrin, Ulrich, Hutchinson, and Norman, 1964) attack was observed for a major part of the duration of the shock delivery, being interrupted primarily by the competing motor reactions elicited by the shock. It appears that shock creates a motivational state in which the opportunity to attack is a reinforcing event for the duration of stimulation and for a brief period after the cessation of stimulation.

References

- AZRIN, N. H., HUTCHINSON, R. R., and HAKE, D. F., Pain-induced fighting in the squirrel monkey. *J. exp. Anal. Behav.*, 1963, 6, 620.
- AZRIN, N. H., HUTCHINSON, R. R., and SALLERY, R. D., Pain-aggression toward inanimate objects. *J. exp. Anal. Behav.*, 1964, 7, 223-228.
- AZRIN, N. H., ULRICH, R. E., HUTCHINSON, R. R., and NORMAN, D. G., Effect of shock duration on shock-induced fighting. *J. exp. Anal. Behav.*, 1964, 7, 9-11.
- DINSMOOR, J. A. and HUGHES, L. H., Training rats to press a bar to turn off shock. *J. comp. physiol. Psychol.*, 1956, 49, 235-238.
- HAKE, D. F. and AZRIN, N. H., An apparatus for delivering pain shock to monkeys. *J. exp. Anal. Behav.*, 1963, 6, 297.

- HUTCHINSON, R. R., AZRIN, N. H., and ULRICH, R. E., Effects of age and related factors on reflexive aggression. *J. comp. physiol. Psychol.* (in press).
- OLDS, J. and MILNER, P., Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. comp. physiol. Psychol.*, 1954, **47**, 419-427.
- PETERSON, N., Control of behavior by presentation of an imprinted stimulus. *Science*, 1960, **132**, 1395-1396.
- SCOTT, J. P., *Aggression*. Chicago: The University of Chicago, 1958.
- SIDMAN, M., Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science*, 1953, **118**, 157-158.
- SKINNER, B. F., *The behavior of organisms: an experimental analysis*. New York: Appleton-Century-Crofts, 1938.
- ULRICH, R. E. and AZRIN, N. H., Reflective fighting in response to aversive stimulation. *J. exp. Anal. Behav.*, 1962, **5**, 511-520.

psychosomatic symptoms that are mediated by the autonomic nervous system. Whereas the former are supposed to be subject to a higher type of control that is symbolic, the latter are presumed to be only the direct physiological consequences of the type and intensity of the patient's emotions (see, for example, 4).

Similarly, students of learning have made a distinction between a lower form, called classical conditioning and thought to be involuntary, and a superior form variously called trial-and-error learning, operant conditioning, type II conditioning, or instrumental learning and believed to be responsible for voluntary behavior. In classical conditioning, the reinforcement must be by an unconditioned stimulus that already elicits the specific response to be learned; therefore, the possibilities are quite limited. In instrumental learning, the reinforcement, called a reward, has the property of strengthening any immediately preceding response. Therefore, the possibilities for reinforcement are much greater; a given reward may reinforce any one of a number of different responses, and a given response may be reinforced by any one of a number of different rewards.

Finally, the foregoing invidious distinctions have coalesced into the strong traditional belief that the superior type of instrumental learning involved in the superior voluntary behavior is possible only for skeletal responses mediated by the superior cerebrospinal nervous system, while, conversely, the inferior classical conditioning is the only kind possible for the inferior, presumably involuntary, visceral and emotional responses mediated by the inferior autonomic nervous system. Thus, in a recent

summary generally considered authoritative, Kimble (5) states the almost universal belief that "for autonomically mediated behavior, the evidence points unequivocally to the conclusion that such responses can be modified by classical, but not instrumental, training methods." Upon examining the evidence, however, one finds that it consists only of failure to secure instrumental learning in two incompletely reported exploratory experiments and a vague allusion to the Russian literature (6). It is only against a cultural background of great prejudice that such weak evidence could lead to such a strong conviction.

The belief that instrumental learning is possible only for the cerebrospinal system and, conversely, that the autonomic nervous system can be modified only by classical conditioning has been used as one of the strongest arguments for the notion that instrumental learning and classical conditioning are two basically different phenomena rather than different manifestations of the same phenomenon under different conditions. But for many years I have been impressed with the similarity between the laws of classical conditioning and those of instrumental learning, and with the fact that, in each of these two situations, some of the specific details of learning vary with the specific conditions of learning. Failing to see any clear-cut dichotomy, I have assumed that there is only one kind of learning (7). This assumption has logically demanded that instrumental training procedures be able to produce the learning of any visceral responses that could be acquired through classical conditioning procedures. Yet it was only a little over a dozen years ago that I began some ex-

perimental work on this problem and a somewhat shorter time ago that I first, in published articles (8), made specific sharp challenges to the traditional view that the instrumental learning of visceral responses is impossible.

Some Difficulties

One of the difficulties of investigating the instrumental learning of visceral responses stems from the fact that the responses that are the easiest to measure—namely, heart rate, vasomotor responses, and the galvanic skin response—are known to be affected by skeletal responses, such as exercise, breathing, and even tensing of certain muscles, such as those in the diaphragm. Thus, it is hard to rule out the possibility that, instead of directly learning a visceral response, the subject has learned a skeletal response the performance of which causes the visceral change being recorded.

One of the controls I planned to use was the paralysis of all skeletal responses through administration of curare, a drug which selectively blocks the motor end plates of skeletal muscles without eliminating consciousness in human subjects or the neural control of visceral responses, such as the beating of the heart. The muscles involved in breathing are paralyzed, so the subject's breathing must be maintained through artificial respiration. Since it seemed unlikely that curarization and other rigorous control techniques would be easy to use with human subjects, I decided to concentrate first on experiments with animals.

Originally I thought that learning would be more difficult when the animal was paralyzed, under the influ-

ence of curare, and therefore I decided to postpone such experiments until ones on nonparalyzed animals had yielded some definitely promising results. This turned out to be a mistake because, as I found out much later, paralyzing the animal with curare not only greatly simplifies the problem of recording visceral responses without artifacts introduced by movement but also apparently makes it easier for the animal to learn, perhaps because paralysis of the skeletal muscles removes sources of variability and distraction. Also, in certain experiments I made the mistake of using rewards that induced strong unconditioned responses that interfered with instrumental learning.

One of the greatest difficulties, however, was the strength of the belief that instrumental learning of glandular and visceral responses is impossible. It was extremely difficult to get students to work on this problem, and when paid assistants were assigned to it, their attempts were so half-hearted that it soon became more economical to let them work on some other problem which they could attack with greater faith and enthusiasm. These difficulties and a few preliminary encouraging but inconclusive early results have been described elsewhere (9).

Success with Salivation

The first clear-cut results were secured by Alfredo Carmona and me in an experiment on the salivation of dogs. Initial attempts to use food as a reward for hungry dogs were unsuccessful, partly because of strong and persistent unconditioned salivation elicited by the food. Therefore, we decided to use water as a reward for thirsty dogs. Preliminary observa-

tions showed that the water had no appreciable effects one way or the other on the bursts of spontaneous salivation. As an additional precaution, however, we used the experimental design of rewarding dogs in one group whenever they showed a burst of spontaneous salivation, so that they would be trained to increase salivation, and rewarding dogs in another group whenever there was a long interval between spontaneous bursts, so that they would be trained to decrease salivation. If the reward had any unconditioned effect, this effect might be classically conditioned to the experimental situation and therefore produce a change in salivation that was not a true instance of instrumental learning. But in classical conditioning the reinforcement must elicit the response that is to be acquired. Therefore, conditioning of a response elicited by the reward could produce either an increase or a decrease in salivation, depending upon the direction of the unconditioned response elicited by the reward, but it could not produce a change in one direction for one group and in the opposite direction for the other group. The same type of logic applies for any unlearned cumulative aftereffects of the reward; they could not be in opposite directions for the two groups. With instrumental learning, however, the reward can reinforce any response that immediately precedes it; therefore, the same reward can be used to produce either increases or decreases.

The results are presented in Fig. 1, which summarizes the effects of 40 days of training with one 45-minute training session per day. It may be seen that in this experiment the learning proceeded slowly. However, statistical analysis showed that each of

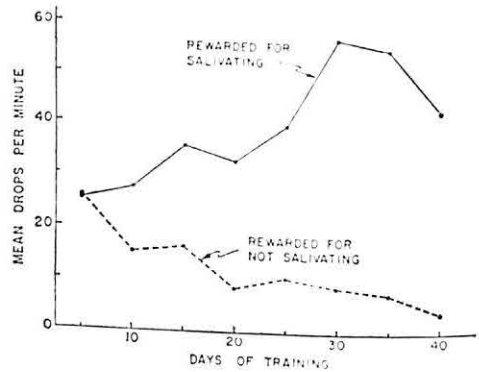


Figure 1
Learning curves for groups of thirsty dogs rewarded with water for either increases or decreases in spontaneous salivation. [From Miller and Carmona (10)]

the trends in the predicted rewarded direction was highly reliable (10).

Since the changes in salivation for the two groups were in opposite directions, they cannot be attributed to classical conditioning. It was noted, however, that the group rewarded for increases seemed to be more aroused and active than the one rewarded for decreases. Conceivably, all we were doing was to change the level of activation of the dogs, and this change was, in turn, affecting the salivation. Although we did not observe any specific skeletal responses, such as chewing movements or panting, which might be expected to elicit salivation, it was difficult to be absolutely certain that such movements did not occur. Therefore, we decided to rule out such movements by paralyzing the dogs with curare, but we immediately found that curare had two effects which were disastrous for this experiment: it elicited such copious and continuous salivation that there were no changes in salivation to reward,

and the salivation was so viscous that it almost immediately gummed up the recording apparatus.

Heart Rate

In the meantime, Jay Trowill, working with me on this problem, was displaying great ingenuity, courage, and persistence in trying to produce instrumental learning of heart rate in rats that had been paralyzed by curare to prevent them from "cheating" by muscular exertion to speed up the heart or by relaxation to slow it down. As a result of preliminary testing, he selected a dose of curare (3.6 milligrams of *d*-tubocurarine chloride per kilogram, injected intraperitoneally) which produced deep paralysis for at least 3 hours, and a rate of artificial respiration (inspiration-expiration ratio 1:1; 70 breaths per minute; peak pressure reading, 20 cm-H₂O) which maintained the heart at a constant and normal rate throughout this time.

In subsequent experiments, DiCara and I have obtained similar effects by starting with a smaller dose (1.2 milligrams per kilogram) and constantly infusing additional amounts of the drug, through intraperitoneal injection, at the rate of 1.2 milligrams per kilogram per hour, for the duration of the experiment. We have recorded, electromyographically, the response of the muscles, to determine that this dose does indeed produce a complete block of the action potentials, lasting for at least an hour after the end of infusion. We have found that if parameters of respiration and the face mask are adjusted carefully, the procedure not only maintains the heart rate of a 500-gram control animal constant but also maintains the

vital signs of temperature, peripheral vasomotor responses, and the $p\text{CO}_2$ of the blood constant.

Since there are not very many ways to reward an animal completely paralyzed by curare, Trowill and I decided to use direct electrical stimulation of rewarding areas of the brain. There were other technical difficulties to overcome, such as devising the automatic system for rewarding small changes in heart rate as recorded by the electrocardiogram. Nevertheless, Trowill at last succeeded in training his rats (11). Those rewarded for an increase in heart rate showed a statistically reliable increase, and those rewarded for a decrease in heart rate showed a statistically reliable decrease. The changes, however, were disappointingly small, averaging only 5 percent in each direction.

The next question was whether larger changes could be achieved by improving the technique of training. DiCara and I used the technique of shaping—in other words, of immediately rewarding first very small, and hence frequently occurring, changes in the correct direction and, as soon as these had been learned, requiring progressively larger changes as the criterion for reward. In this way, we were able to produce in 90 minutes of training changes averaging 20 percent in either direction (12).

Key Properties of Learning: Discrimination and Retention

Does the learning of visceral responses have the same properties as the learning of skeletal responses? One of the important characteristics of the instrumental learning of skeletal responses is that a discrimination can be learned, so that the responses

are more likely to be made in the stimulus situations in which they are rewarded than in those in which they are not. After the training of the first few rats had convinced us that we could produce large changes in heart rate, DiCara and I gave all the rest of the rats in the experiment described above 45 minutes of additional training with the most difficult criterion. We did this in order to see whether they could learn to give a greater response during a "time-in" stimulus (the presence of a flashing light and a tone) which indicated that a response in the proper direction would be rewarded than during a "time-out" stimulus (absence of light and tone) which indicated that a correct response would not be rewarded.

Figure 2 shows the record of one of the rats given such training. Before the beginning of the special discrimination training it had slowed its heart

from an initial rate of 350 beats per minute to a rate of 230 beats per minute. From the top record of Fig. 2 one can see that, at the beginning of the special discrimination training, there was no appreciable reduction in heart rate that was specifically associated with the time-in stimulus. Thus it took the rat considerable time after the onset of this stimulus to meet the criterion and get the reward. At the end of the discrimination training the heart rate during time-out remained approximately the same, but when the time-in light and tone came on, the heart slowed down and the criterion was promptly met. Although the other rats showed less change than this, by the end of the relatively short period of discrimination training their heart rate did change reliably ($P < .001$) in the predicted direction when the time-in stimulus came on. Thus, it is clear that instrumental visceral learn-

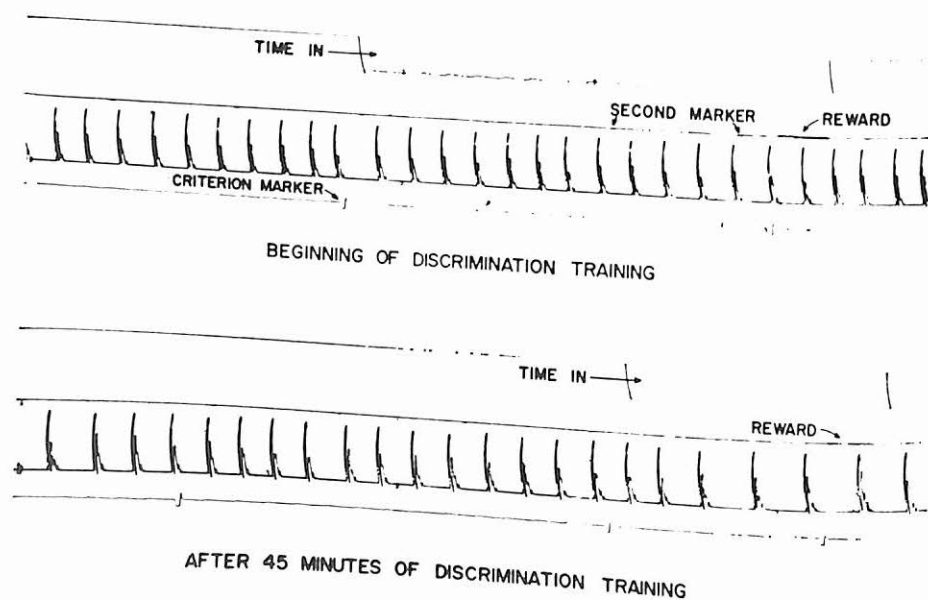


Figure 2
Electrocardiograms at the beginning and at the end of discrimination training of curarized rat rewarded for slow heart rate. Slowing of heart rate is rewarded only during a "time-in" stimulus (tone and light). [From Miller and DiCara (12)]

ing has at least one of the important properties of instrumental skeletal learning—namely, the ability to be brought under the control of a discriminative stimulus.

Another of the important properties of the instrumental learning of skeletal responses is that it is remembered. DiCara and I performed a special experiment to test the retention of learned changes in heart rate (13). Rats that had been given a single training session were returned to their home cages for 3 months without further training. When curarized again and returned to the experimental situation for nonreinforced test trials, rats in both the "increase" and the "decrease" groups showed good retention by exhibiting reliable changes in the direction rewarded in the earlier training.

Escape and Avoidance Learning

Is visceral learning by any chance peculiarly limited to reinforcement by the unusual reward of direct electrical stimulation of the brain, or can it be reinforced by other rewards in the same way that skeletal learning can be? In order to answer this question, DiCara and I (14) performed an experiment using the other of the two forms of thoroughly studied reward that can be conveniently used with rats which are paralyzed by curare—namely, the chance to avoid, or escape from, mild electric shock. A shock signal was turned on; after it had been on for 10 seconds it was accompanied by brief pulses of mild electric shock delivered to the rat's tail. During the first 10 seconds the rat could turn off the shock signal and avoid the shock by making the correct response of changing its heart rate in the required direction by the required

amount. If it did not make the correct response in time, the shocks continued to be delivered until the rat escaped them by making the correct response, which immediately turned off both the shock and the shock signal.

For one group of curarized rats, the correct response was an increase in heart rate; for the other group it was a decrease. After the rats had learned to make small responses in the proper direction, they were required to make larger ones. During this training the shock signals were randomly interspersed with an equal number of "safe" signals that were not followed by shock; the heart rate was also recorded during so-called blank trials—trials without any signals or shocks. For half of the rats the shock signal was a tone and the "safe" signal was a flashing light; for the other half the roles of these cues were reversed.

The results are shown in Fig. 3. Each of the 12 rats in this experiment changed its heart rate in the rewarded direction. As training progressed, the shock signal began to elicit a progressively greater change in the rewarded direction than the change recorded

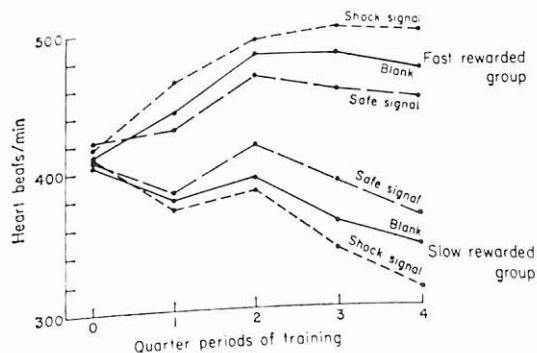


Figure 3

Changes in heart rate during avoidance training. [From DiCara and Miller (14)]

during the blank trials; this was a statistically reliable trend. Conversely, as training progressed, the "safe" signal came to elicit a statistically reliable change in the opposite direction, toward the initial base line. These results show learning when escape and avoidance are the rewards; this means that visceral responses in curarized rats can be reinforced by rewards other than direct electrical stimulation of the brain. These rats also discriminate between the shock and the "safe" signals. You will remember that, with noncurarized thirsty dogs, we were able to use yet another kind of reward, water, to produce learned changes in salivation.

Transfer to Noncurarized State: More Evidence Against Mediation

In the experiments discussed above, paralysis of the skeletal muscles by curare ruled out the possibility that the subjects were learning the overt performance of skeletal responses which were indirectly eliciting the changes in the heart rate. It is barely conceivable, however, that the rats were learning to send out from the motor cortex central impulses which would have activated the muscles had they not been paralyzed. And it is barely conceivable that these central impulses affected heart rate by means either of inborn connections or of classically conditioned ones that had been acquired when previous exercise had been accompanied by an increase in heart rate and relaxation had been accompanied by a decrease. But, if the changes in heart rate were produced in this indirect way, we would expect that, during a subsequent test without curare, any rat that showed learned changes in heart rate would show the movements in the muscles that were

no longer paralyzed. Furthermore, the problem of whether or not visceral responses learned under curarization carry over to the noncurarized state is of interest in its own right.

In order to answer this question, DiCara and I (15) trained two groups of curarized rats to increase or decrease, respectively, their heart rate in order to avoid, or escape from, brief pulses of mild electric shock. When these rats were tested 2 weeks later in the noncurarized state, the habit was remembered. Statistically reliable increases in heart rate averaging 5 percent and decreases averaging 16 percent occurred. Immediately subsequent retraining without curare produced additional significant changes of heart rate in the rewarded direction, bringing the total overall increase to 11 percent and the decrease to 22 percent. While at the beginning of the test in the noncurarized state, the two groups showed some differences in respiration and activity, these differences decreased until, by the end of the retraining, they were small and far from statistically reliable ($t = 0.3$ and 1.3 , respectively). At the same time, the difference between the two groups with respect to heart rate was increasing, until it became large and thus extremely reliable ($t = 8.6$, d.f. = 12, $P < .001$).

In short, while greater changes in heart rate were being learned, the response was becoming more specific, involving smaller changes in respiration and muscular activity. This increase in specificity with additional training is another point of similarity with the instrumental learning of skeletal responses. Early in skeletal learning, the rewarded correct response is likely to be accompanied by many unnecessary movements. With additional training during which ex-

traneous movements are not rewarded, they tend to drop out.

It is difficult to reconcile the foregoing results with the hypothesis that the differences in heart rate were mediated primarily by a difference in either respiration or amount of general activity. This is especially true in view of the research, summarized by Ehrlich and Malmo (16), which shows that muscular activity, to affect heart rate in the rat, must be rather vigorous.

While it is difficult to rule out completely the possibility that changes in heart rate are mediated by central impulses to skeletal muscles, the possibility of such mediation is much less attractive for other responses, such as intestinal contractions and the formation of urine by the kidney. Furthermore, if the learning of these different responses can be shown to be specific in enough visceral responses, one runs out of different skeletal movements each eliciting a specific different visceral response (17). Therefore, experiments were performed on the learning of a variety of different visceral responses and on the specificity of that learning. Each of these experiments was, of course, interesting in its own right, quite apart from any bearing on the problem of mediation.

Specificity: Intestinal Versus Cardiac

The purpose of our next experiment was to determine the specificity of visceral learning. If such learning has the same properties as the instrumental learning of skeletal responses, it should be possible to learn a specific visceral response independently of other ones. Furthermore, as we have just seen, we might expect to find that, the better the rewarded response is learned, the more specific is the

learning. Banuazizi and I worked on this problem (18). First we had to discover another visceral response that could be conveniently recorded and rewarded. We decided on intestinal contractions, and recorded them in the curarized rat with a little balloon filled with water thrust approximately 4 centimeters beyond the anal sphincter. Changes of pressure in the balloon were transduced into electric voltages which produced a record on a polygraph and also activated an automatic mechanism for delivering the reward, which was electrical stimulation of the brain.

The results for the first rat trained, which was a typical one, are shown in Fig. 4. From the top record it may be seen that, during habituation, there were some spontaneous contractions. When the rat was rewarded by brain stimulation for keeping contractions below a certain amplitude for a certain time, the number of contractions was reduced and the base line was lowered. After the record showed a highly reliable change indicating that relaxation had been learned (Fig. 4, second record from the top) the conditions of training were reversed and the reward was delivered whenever the amplitude of contractions rose above a certain level. From the next record (Fig. 4, middle) it may be seen that this type of training increased the number of contractions and raised the base line. Finally (Fig. 4, two bottom records) the reward was discontinued and, as would be expected, the response continued for a while but gradually became extinguished, so that the activity eventually returned to approximately its original base-line level.

After studying a number of other rats in this way and convincing ourselves that the instrumental learning

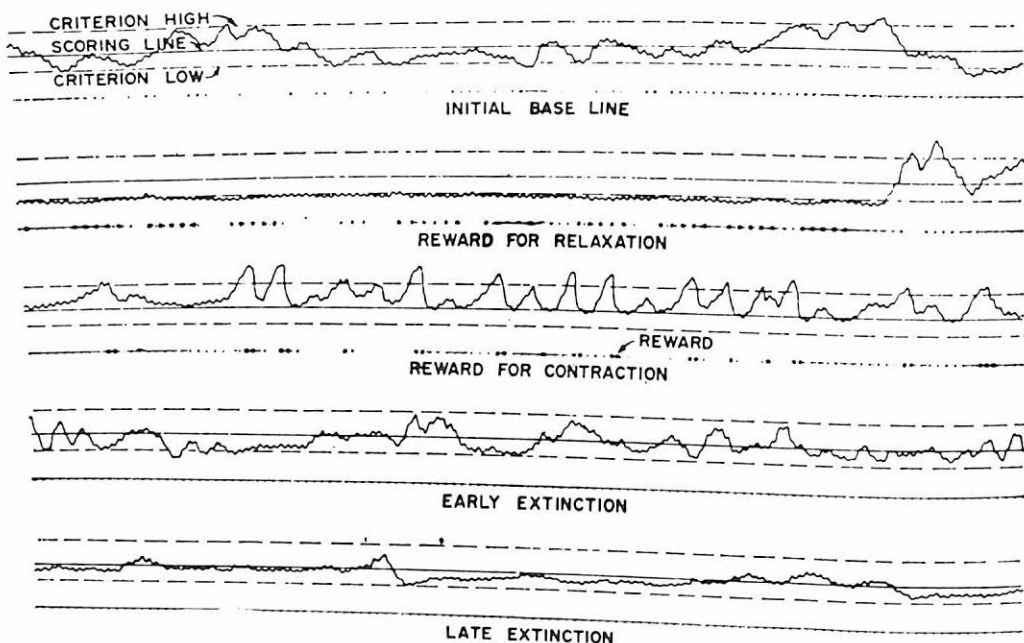


Figure 4
Typical samples of a record of instrumental learning of an intestinal response by a curarized rat. (From top to bottom) Record of spontaneous contraction before training; record after training with reward for relaxation; record after training with reward for contractions; records during nonrewarded extinction trials. [From Miller and Banuazizi (18)]

of intestinal responses was a possibility, we designed an experiment to test specificity. For all the rats of the experiment, both intestinal contractions and heart rate were recorded, but half the rats were rewarded for one of these responses and half were rewarded for the other response. Each of these two groups of rats was divided into two subgroups, rewarded, respectively, for increased and decreased response. The rats were completely paralyzed by curare, maintained on artificial respiration, and rewarded by electrical stimulation of the brain.

The results are shown in Figs. 5 and 6. In Fig. 5 it may be seen that

the group rewarded for increases in intestinal contractions learned an increase, the group rewarded for decreases learned a decrease, but neither of these groups showed an appreciable change in heart rate. Conversely (Fig. 6), the group rewarded for increases in heart rate showed an increase, the group rewarded for decreases showed a decrease, but neither of these groups showed a change in intestinal contractions.

The fact that each type of response changed when it was rewarded rules out the interpretation that the failure to secure a change when that change was not rewarded could have been due to either a strong and stable homeo-

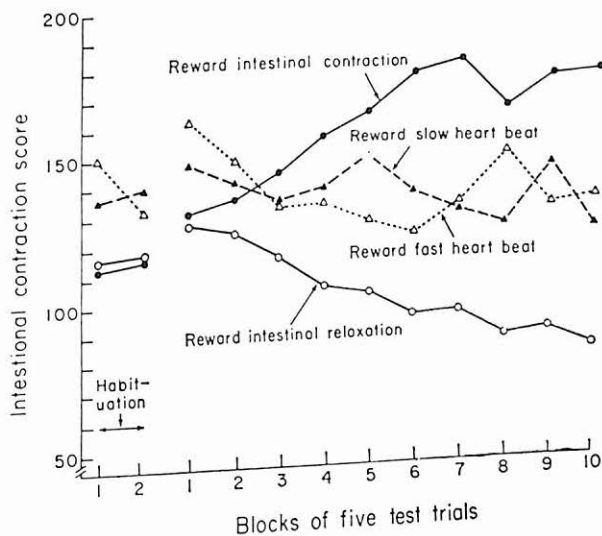


Figure 5

Graph showing that the intestinal contraction score is changed by rewarding either increases or decreases in intestinal contractions but is unaffected by rewarding changes in heart rate. [From Miller and Banuazizi (18)]

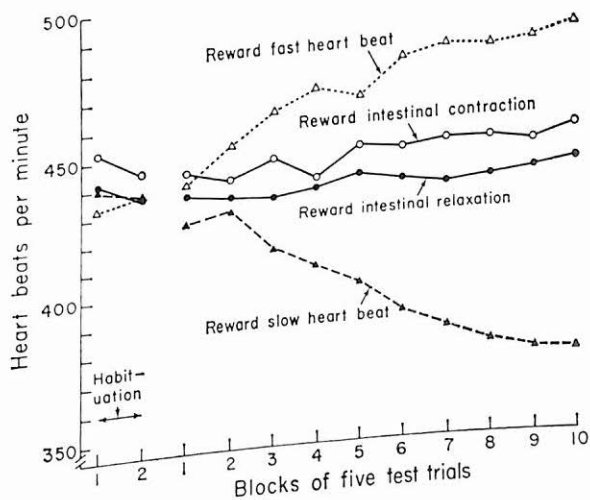


Figure 6

Graph showing that the heart rate is changed by rewarding either increases or decreases in heart rate but is unaffected by rewarding changes in visceral contractions. Comparison with Fig. 5 demonstrates the specificity of visceral learning. [From Miller and Banuazizi (18)]

static regulation of that response or an inability of our techniques to measure changes reliably under the particular conditions of our experiment.

Each of the 12 rats in the experiment showed statistically reliable changes in the rewarded direction; for 11 the changes were reliable beyond the $P < .001$ level, while for the 12th the changes were reliable only beyond the .05 level. A statistically reliable negative correlation showed that the better the rewarded visceral response was learned, the less change occurred in the other, nonrewarded response. This greater specificity with better learning is what we had expected. The results showed that visceral learning can be specific to an organ system, and they clearly ruled out the possibility of mediation by any single general factor, such as level of activation or central commands for either general activity or relaxation.

In an additional experiment, Banuazizi (19) showed that either increases or decreases in intestinal contractions can be rewarded by avoidance of, or escape from, mild electric shocks, and that the intestinal responses can be discriminatively elicited by a specific stimulus associated with reinforcement.

Kidney Function

Encouraged by these successes, Di-Cara and I decided to see whether or not the rate of urine formation by the kidney could be changed in the curarized rat rewarded by electrical stimulation of the brain (20). A catheter, permanently inserted, was used to prevent accumulation of urine by the bladder, and the rate of urine formation was measured by an electronic

device for counting minute drops. In order to secure a rate of urine formation fast enough so that small changes could be promptly detected and rewarded, the rats were kept constantly loaded with water through infusion by way of a catheter permanently inserted in the jugular vein.

All of the seven rats rewarded when the intervals between times of urine-drop formation lengthened showed decreases in the rate of urine formation, and all of the seven rats rewarded when these intervals shortened showed increases in the rate of urine formation. For both groups the changes were highly reliable ($P < .001$).

In order to determine how the change in rate of urine formation was achieved, certain additional measures were taken. As the set of bars at left in Fig. 7 shows, the rate of filtration, measured by means of ^{14}C -labeled inulin, increased when increases in the rate of urine formation were rewarded and decreased when decreases in the rate were rewarded. Plots of the correlations showed that the changes in the rates of filtration and urine formation were not related to changes in either blood pressure or heart rate.

The middle set of bars in Fig. 7 shows that the rats rewarded for increases in the rate of urine formation had an increased rate of renal blood flow, as measured by ^3H -*p*-aminohippuric acid, and that those rewarded for decreases had a decreased rate of renal blood flow. Since these changes in blood flow were not accompanied by changes in general blood pressure or in heart rate, they must have been achieved by vasomotor changes of the renal arteries. That these vasomotor changes were at least somewhat specific is shown by the fact that vasomotor responses of the tail, as measured by

a photoelectric plethysmograph, did not differ for the two groups of rats.

The set of bars at right in Fig. 7 shows that when decreases in rate of urine formation were rewarded, a more concentrated urine, having higher osmolarity, was formed. Since the slower passage of urine through the tubules would afford more opportunity for reabsorption of water, this higher concentration does not necessarily mean an increase in the secretion of antidiuretic hormone. When an increased rate of urine formation was rewarded, the urine did not become more diluted—that is, it showed no decrease in osmolarity; therefore, the increase in rate of urine formation observed in this experiment cannot be accounted for in terms of an inhibition of the secretion of antidiuretic hormone.

From the foregoing results it appears that the learned changes in urine formation in this experiment were produced primarily by changes in the rate of filtration, which, in turn, were produced primarily by changes in the rate of blood flow through the kidneys.

Gastric Changes

In the next experiment, Carmona, Demierre, and I used a photoelectric plethysmograph to measure changes, presumably in the amount of blood, in the stomach wall (21). In an operation performed under anesthesia, a small glass tube, painted black except for a small spot, was inserted into the rat's stomach. The same tube was used to hold the stomach wall against a small glass window inserted through the body wall. The tube was left in that position. After the animal had recovered, a bundle of optical fibers could be slipped snugly into the glass tube so that the light beamed through it would shine out through the unpainted spot in the tube inside the stomach, pass through the stomach wall, and be recorded by a photocell on the other side of the glass window. Preliminary tests indicated that, as would be expected, when the amount of blood in the stomach wall increased, less light would pass through. Other tests showed that stomach contractions elicited by injections of in-

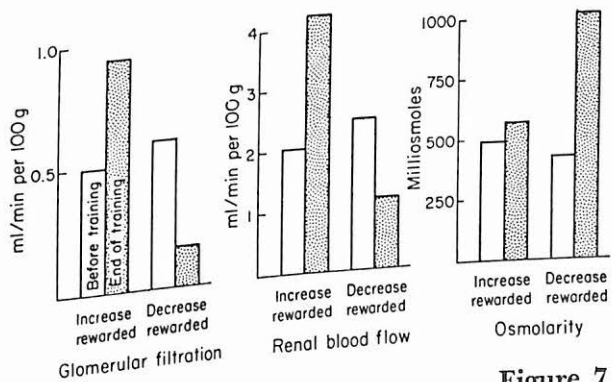


Figure 7

Effects of rewarding increased rate of urine formation in one group and decreased rate in another on measures of glomerular filtration, renal blood flow, and osmolarity. [From data in Miller and DiCara (20)]

sulin did not affect the amount of light transmitted.

In the main experiment we rewarded curarized rats by enabling them to avoid or escape from mild electric shocks. Some were rewarded when the amount of light that passed through the stomach wall increased, while others were rewarded when the amount decreased. Fourteen of the 15 rats showed changes in the rewarded direction. Thus, we demonstrated that the stomach wall, under the control of the autonomic nervous system, can be modified by instrumental learning. There is strong reason to believe that the learned changes were achieved by vasomotor responses affecting the amount of blood in the stomach wall or mucosa, or in both.

In another experiment, Carmona (22) showed that stomach contractions can be either increased or decreased by instrumental learning.

It is obvious that learned changes in the blood supply of internal organs can affect their functioning—as, for example, the rate at which urine was formed by the kidneys was affected by changes in the amount of blood that flowed through them. Thus, such changes can produce psychosomatic symptoms. And if the learned changes in blood supply can be specific to a given organ, the symptom will occur in that organ rather than in another one.

Peripheral Vasomotor Responses

Having investigated the instrumental learning of internal vasomotor responses, we next studied the learning of peripheral ones. In the first experiment, the amount of blood in the tail of a curarized rat was measured by a photoelectric plethysmograph, and changes were rewarded by electrical

stimulation of the brain (23). All of the four rats rewarded for vasoconstriction showed that response, and, at the same time, their average core temperature, measured rectally, decreased from 98.9° to 97.9°F. All of the four rats rewarded for vasodilatation showed that response and, at the same time, their average core temperature increased from 99.9° to 101°F. The vasomotor change for each individual rat was reliable beyond the $P < .01$ level, and the difference in change in temperature between the groups was reliable beyond the .01 level. The direction of the change in temperature was opposite to that which would be expected from the heat conservation caused by peripheral vasoconstriction or the heat loss caused by peripheral vasodilatation. The changes are in the direction which would be expected if the training had altered the rate of heat production, causing a change in temperature which, in turn, elicited the vasomotor response.

The next experiment was designed to try to determine the limits of the specificity of vasomotor learning. The pinnae of the rat's ears were chosen because the blood vessels in them are believed to be innervated primarily, and perhaps exclusively, by the sympathetic branch of the autonomic nervous system, the branch that Cannon believed always fired nonspecifically as a unit (3). But Cannon's experiments involved exposing cats to extremely strong emotion-evoking stimuli, such as barking dogs, and such stimuli will also evoke generalized activity throughout the skeletal musculature. Perhaps his results reflected the way in which sympathetic activity was elicited, rather than demonstrating any inherent inferiority of the sympathetic nervous system.

In order to test this interpretation,

DiCara and I (24) put photocells on both ears of the curarized rat and connected them to a bridge circuit so that only differences in the vasomotor responses of the two ears were rewarded by brain stimulation. We were somewhat surprised and greatly delighted to find that this experiment actually worked. The results are summarized in Fig. 8. Each of the six rats rewarded for relative vasodilatation of the left ear showed that response, while each of the six rats rewarded for relative vasodilatation of the right ear showed that response. Recordings from the right and left forepaws showed little if any change in vasomotor response.

It is clear that these results cannot be by-products of changes in either heart rate or blood pressure, as these would be expected to affect both ears equally. They show either that vasomotor responses mediated by the sympathetic nervous system are capable

of much greater specificity than has previously been believed, or that the innervation of the blood vessels in the pinnae of the ears is not restricted almost exclusively to sympathetic-nervous-system components, as has been believed, and involves functionally significant parasympathetic components. In any event, the changes in the blood flow certainly were surprisingly specific. Such changes in blood flow could account for specific psychosomatic symptoms.

Blood Pressure Independent of Heart Rate

Although changes in blood pressure were not induced as by-products of rewarded changes in the rate of urine formation, another experiment on curarized rats showed that, when changes in systolic blood pressure are specifically reinforced, they can be learned (25). Blood pressure was re-

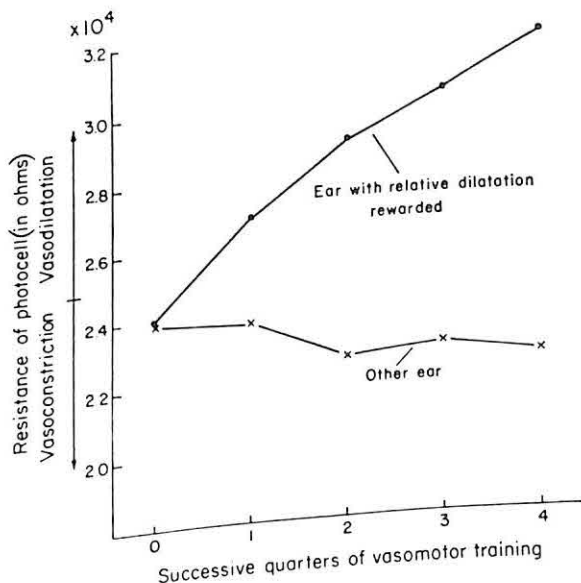


Figure 8

Learning a difference in the vasomotor responses of the two ears in the curarized rat. [From data in DiCara and Miller (24)]

corded by means of a catheter permanently inserted into the aorta, and the reward was avoidance of, or escape from, mild electric shock. All seven rats rewarded for increases in blood pressure showed further increases, while all seven rewarded for decreases showed decreases, each of the changes, which were in opposite directions, being reliable beyond the $P < .01$ level. The increase was from 139 mm-Hg, which happens to be roughly comparable to the normal systolic blood pressure of an adult man, to 170 mm-Hg, which is on the borderline of abnormally high blood pressure in man.

Each experimental animal was "yoked" with a curarized partner, maintained on artificial respiration and having shock electrodes on its tail wired in series with electrodes on the tail of the experimental animal, so that it received exactly the same electric shocks and could do nothing to escape or avoid them. The yoked controls for both the increase-rewarded and the decrease-rewarded groups showed some elevation in blood pressure as an unconditioned effect of the shocks. By the end of training, in contrast to the large difference in the blood pressures of the two groups specifically rewarded for changes in opposite directions, there was no difference in blood pressure between the yoked control partners for these two groups. Furthermore, the increase in blood pressure in these control groups was reliably less ($P < .01$) than that in the group specifically rewarded for increases. Thus, it is clear that the reward for an increase in blood pressure produced an additional increase over and above the effects of the shocks per se, while the reward for a decrease was able to overcome the unconditioned increase elicited by the shocks.

For none of the four groups was there a significant change in heart rate or in temperature during training; there were no significant differences in these measures among the groups. Thus, the learned change was relatively specific to blood pressure.

Transfer from Heart Rate to Skeletal Avoidance

Although visceral learning can be quite specific, especially if only a specific response is rewarded, as was the case in the experiment on the two ears, under some circumstances it can involve a more generalized effect.

In handling the rats that had just recovered from curarization, DiCara noticed that those that had been trained, through the avoidance or escape reward, to increase their heart rate were more likely to squirm, squeal, defecate, and show other responses indicating emotionality than were those that had been trained to reduce their heart rate. Could instrumental learning of heart-rate changes have some generalized effects, perhaps on the level of emotionality, which might affect the behavior in a different avoidance-learning situation? In order to look for such an effect, DiCara and Weiss (26) used a modified shuttle avoidance apparatus. In this apparatus, when a danger signal is given, the rat must run from compartment A to compartment B. If he runs fast enough, he avoids the shock; if not, he must run to escape it. The next time the danger signal is given, the rat must run in the opposite direction, from B to A.

Other work had shown that learning in this apparatus is an inverted U-shaped function of the strength of the shocks, with shocks that are too strong eliciting emotional behavior instead of running. DiCara and Weiss trained

their rats in this apparatus with a level of shock that is approximately optimum for naive rats of this strain. They found that the rats that had been rewarded for decreasing their heart rate learned well, but that those that had been rewarded for increasing their heart rate learned less well, as if their emotionality had been increased. The difference was statistically reliable ($P < .001$). This experiment clearly demonstrates that training a visceral response can affect the subsequent learning of a skeletal one, but additional work will be required to prove the hypothesis that training to increase heart rate increases emotionality.

Visceral Learning Without Curare

Thus far, in all of the experiments except the one on teaching thirsty dogs to salivate, the initial training was given when the animal was under the influence of curare. All of the experiments, except the one on salivation, have produced surprisingly rapid learning—definitive results within 1 or 2 hours. Will learning in the normal, noncurarized state be easier, as we originally thought it should be, or will it be harder, as the experiment on the noncurarized dogs suggests? DiCara and I have started to get additional evidence on this problem. We have obtained clear-cut evidence that rewarding (with the avoidance or escape reward) one group of freely moving rats for reducing heart rate and rewarding another group for increasing heart rate produces a difference between the two groups (27). That this difference was not due to the indirect effects of the overt performance of skeletal responses is shown by the fact that it persisted in subsequent tests during which the rats were paralyzed by curare. And,

on subsequent retraining without curare, such differences in activity and respiration as were present earlier in training continued to decrease, while the differences in heart rate continued to increase. It seems extremely unlikely that, at the end of training, the highly reliable differences in heart rate ($t = 7.2$; $P < .0001$) can be explained by the highly unreliable differences in activity and respiration ($t = .07$ and 0.2 , respectively).

Although the rats in this experiment showed some learning when they were trained initially in the noncurarized state, this learning was much poorer than that which we have seen in our other experiments on curarized rats. This is exactly the opposite of my original expectation, but seems plausible in the light of hindsight. My hunch is that paralysis by curare improved learning by eliminating sources of distraction and variability. The stimulus situation was kept more constant, and confusing visceral fluctuations induced indirectly by skeletal movements were eliminated.

Learned Changes in Brain Waves

Encouraged by success in the experiments on the instrumental learning of visceral responses, my colleagues and I have attempted to produce other unconventional types of learning. Electrodes placed on the skull or, better yet, touching the surface of the brain record summative effects of electrical activity over a considerable area of the brain. Such electrical effects are called brain waves, and the record of them is called an electroencephalogram. When the animal is aroused, the electroencephalogram consists of fast, low-voltage activity; when the animal is drowsy or sleeping normally, the electroencephalo-

logram consists of considerably slower, higher-voltage activity. Carmona attempted to see whether this type of brain activity, and the state of arousal accompanying it, can be modified by direct reward of changes in the brain activity (28, 29).

The subjects of the first experiment were freely moving cats. In order to have a reward that was under complete control and that did not require the cat to move, Carmona used direct electrical stimulation of the medial forebrain bundle, which is a rewarding area of the brain. Such stimulation produced a slight lowering in the average voltage of the electroencephalogram and an increase in behavioral arousal. In order to provide a control for these and any other unlearned effects, he rewarded one group for changes in the direction of high-voltage activity and another group for changes in the direction of low-voltage activity.

Both groups learned. The cats rewarded for high-voltage activity showed more high-voltage slow waves and tended to sit like sphinxes, staring out into space. The cats rewarded for low-voltage activity showed much more low-voltage fast activity, and appeared to be aroused, pacing restlessly about, sniffing, and looking here and there. It was clear that this type of training had modified both the character of the electrical brain waves and the general level of the behavioral activity. It was not clear, however, whether the level of arousal of the brain was directly modified and hence modified the behavior; whether the animals learned specific items of behavior which, in turn, modified the arousal of the brain as reflected in the electroencephalogram; or whether both types of learning were occurring simultaneously.

In order to rule out the direct sensory consequences of changes in muscular tension, movement, and posture, Carmona performed the next experiment on rats that had been paralyzed by means of curare. The results, given in Fig. 9, show that both rewarded groups showed changes in the rewarded direction; that a subsequent nonrewarded rest increased the number of high-voltage responses in both groups; and that, when the conditions of reward were reversed, the direction of change in voltage was reversed.

At present we are trying to use similar techniques to modify the functions of a specific part of the vagal nucleus, by recording and specifically rewarding changes in the electrical activity there. Preliminary results suggest that this is possible. The next step is to investigate the visceral consequences of such modification. This kind of work may open up possibilities for modifying the activity of specific parts of the brain and the functions that they control. In some cases, directly rewarding brain activity may be a more convenient or more powerful technique than rewarding skeletal or visceral behavior. It also may be a new way to throw light on the functions of specific parts of the brain (30).

Human Visceral Learning

Another question is that of whether people are capable of instrumental learning of visceral responses. I believe that in this respect they are as smart as rats. But, as a recent critical review by Katkin and Murray (31) points out, this has not yet been completely proved. These authors have comprehensively summarized the recent studies reporting successful use of instrumental training to modify human heart rate, vasomotor responses,

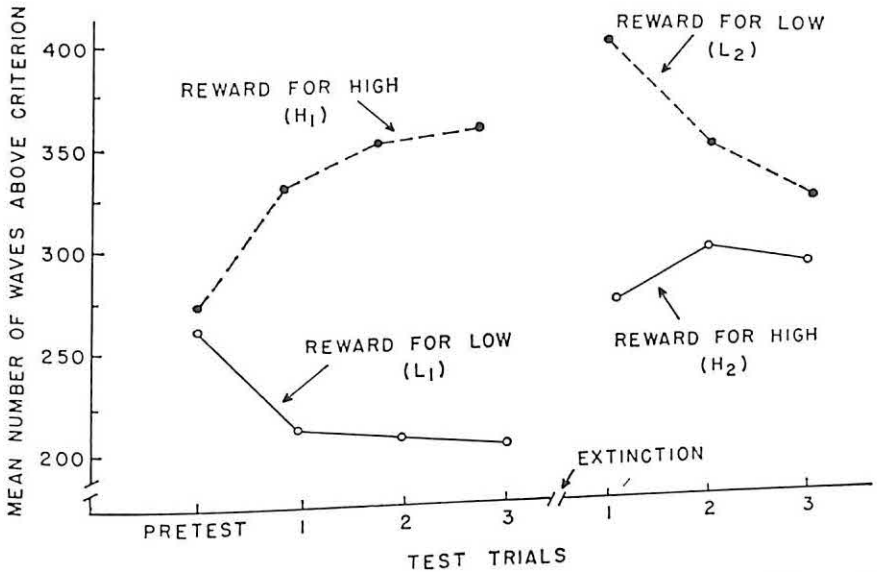


Figure 9

Instrumental learning by curarized rats rewarded for high-voltage or for low-voltage electroencephalograms recorded from the cerebral cortex. After a period of nonrewarded extinction, which produced some drowsiness, as indicated by an increase in voltage, the rats in the two groups were then rewarded for voltage changes opposite in direction to the changes for which they were rewarded earlier. [From Carmona (29)]

and the galvanic skin response. Because of the difficulties in subjecting human subjects to the same rigorous controls, including deep paralysis by means of curare, that can be used with animal subjects, one of the most serious questions about the results of the human studies is whether the changes recorded represent the true instrumental learning of visceral responses or the unconscious learning of those skeletal responses that can produce visceral reactions. However, the able investigators who have courageously challenged the strong traditional belief in the inferiority of the autonomic nervous system with experiments at the more difficult but especially significant human level are developing ingenious controls, in-

cluding demonstrations of the specificity of the visceral change, so that their cumulative results are becoming increasingly impressive.

Possible Role in Homeostasis

The functional utility of instrumental learning by the cerebrospinal nervous system under the conditions that existed during mammalian evolution is obvious. The skeletal responses mediated by the cerebrospinal nervous system operate on the external environment, so that there is survival value in the ability to learn responses that bring rewards such as food, water, or escape from pain. The fact that the responses mediated by the autonomic nervous system do not

have such direct action on the external environment was one of the reasons for believing that they are not subject to instrumental learning. Is the learning ability of the autonomic nervous system something that has no normal function other than that of providing my students with subject matter for publications? Is it a mere accidental by-product of the survival value of cerebrospinal learning, or does the instrumental learning of autonomically mediated responses have some adaptive function, such as helping to maintain that constancy of the internal environment called homeostasis?

In order for instrumental learning to function homeostatically, a deviation away from the optimum level will have to function as a drive to motivate learning, and a change toward the optimum level will have to function as a reward to reinforce the learning of the particular visceral response that produced the corrective change.

When a mammal has less than the optimum amount of water in his body, this deficiency serves as a drive of thirst to motivate learning; the overt consummatory response of drinking functions as a reward to reinforce the learning of the particular skeletal responses that were successful in securing the water that restored the optimum level. But is the consummatory response essential? Can restoration of an optimum level by a glandular response function as a reward?

In order to test for the possible rewarding effects of a glandular response, DiCara, Wolf, and I (32) injected albino rats with antidiuretic hormone (ADH) if they chose one arm of a T-maze and with the isotonic saline vehicle if they chose the other,

distinctively different, arm. The ADH permitted water to be reabsorbed in the kidney, so that a smaller volume of more concentrated urine was formed. Thus, for normal rats loaded in advance with H_2O , the ADH interfered with the excess-water excretion required for the restoration of homeostasis, while the control injection of isotonic saline allowed the excess water to be excreted. And, indeed, such rats learned to select the side of the maze that assured them an injection of saline so that their glandular response could restore homeostasis.

Conversely, for rats with diabetes insipidus, loaded in advance with hypertonic NaCl, the homeostatic effects of the same two injections were reversed; the ADH, causing the urine to be more concentrated, helped the rats to get rid of the excess NaCl, while the isotonic saline vehicle did not. And, indeed, a group of rats of this kind learned the opposite choice of selecting the ADH side of the maze. As a further control on the effects of the ADH per se, normal rats which had not been given H_2O or NaCl exhibited no learning. This experiment showed that an excess of either H_2O or NaCl functions as a drive and that the return to the normal concentration produced by the appropriate response of a gland, the kidney, functions as a reward.

When we consider the results of this experiment together with those of our experiments showing that glandular and visceral responses can be instrumentally learned, we will expect the animal to learn those glandular and visceral responses mediated by the central nervous system that promptly restore homeostasis after any considerable deviation. Whether or not this theoretically possible learning has any practical significance will depend on

whether or not the innate homeostatic mechanisms control the levels closely enough to prevent any deviations large enough to function as a drive from occurring. Even if the innate control should be accurate enough to preclude learning in most cases, there remains the intriguing possibility that, when pathology interferes with innate control, visceral learning is available as a supplementary mechanism.

Implications and Speculations

We have seen how the instrumental learning of visceral responses suggests a new possible homeostatic mechanism worthy of further investigation. Such learning also shows that the autonomic nervous system is not as inferior as has been so widely and firmly believed. It removes one of the strongest arguments for the hypothesis that there are two fundamentally different mechanisms of learning, involving different parts of the nervous system.

CAUSE OF PSYCHOSOMATIC SYMPTOMS

Similarly, evidence of the instrumental learning of visceral responses removes the main basis for assuming that the psychosomatic symptoms that involve the autonomic nervous system are fundamentally different from those functional symptoms, such as hysterical ones, that involve the cerebrospinal nervous system. Such evidence allows us to extend to psychosomatic symptoms the type of learning-theory analysis that Dollard and I (7, 33) have applied to other symptoms.

For example, suppose a child is terror-stricken at the thought of going to school in the morning because he

is completely unprepared for an important examination. The strong fear elicits a variety of fluctuating autonomic symptoms, such as a queasy stomach at one time and pallor and faintness at another; at this point his mother, who is particularly concerned about cardiovascular symptoms, says, "You are sick and must stay home." The child feels a great relief from fear, and this reward should reinforce the cardiovascular responses producing pallor and faintness. If such experiences are repeated frequently enough, the child, theoretically, should learn to respond with that kind of symptom. Similarly, another child whose mother ignored the vasomotor responses but was particularly concerned by signs of gastric distress would learn the latter type of symptom. I want to emphasize, however, that we need careful clinical research to determine how frequently, if at all, the social conditions sufficient for such theoretically possible learning of visceral symptoms actually occur. Since a given instrumental response can be reinforced by a considerable variety of rewards, and by one reward on one occasion and a different reward on another, the fact that glandular and visceral responses can be instrumentally learned opens up many new theoretical possibilities for the reinforcement of psychosomatic symptoms.

Furthermore, we do not yet know how severe a psychosomatic effect can be produced by learning. While none of the 40 rats rewarded for speeding up their heart rates have died in the course of training under curarization, 7 of the 40 rats rewarded for slowing down their heart rates have died. This statistically reliable difference ($\chi^2 = 5.6, P < .02$) is highly suggestive, but it could mean that train-

ing to speed up the heart helped the rats resist the stress of curare rather than that the reward for slowing down the heart was strong enough to overcome innate regulatory mechanisms and induce sudden death. In either event the visceral learning had a vital effect. At present, DiCara and I are trying to see whether or not the learning of visceral responses can be carried far enough in the noncurarized animal to produce physical damage. We are also investigating the possibility that there may be a critical period in early infancy during which visceral learning has particularly intense and long-lasting effects.

INDIVIDUAL AND CULTURAL DIFFERENCES

It is possible that, in addition to producing psychosomatic symptoms in extreme cases, visceral learning can account for certain more benign individual and cultural differences. Lacey and Lacey (34) have shown that a given individual may have a tendency, which is stable over a number of years, to respond to a variety of different stresses with the same profile of autonomic responses, while other individuals may have statistically reliable tendencies to respond with different profiles. It now seems possible that differential conditions of learning may account for at least some of these individual differences in patterns of autonomic response.

Conversely, such learning may account also for certain instances in which the same individual responds to the same stress in different ways. For example, a small boy who receives a severe bump in rough-and-tumble play may learn to inhibit the secretion of tears in this situation since his peer group will punish crying by calling it "sissy." But the same small boy may

burst into tears when he gets home to his mother, who will not punish weeping and may even reward tears with sympathy.

Similarly, it seems conceivable that different conditions of reward by a culture different from our own may be responsible for the fact that Homer's adult heroes so often "let the big tears fall." Indeed, a former colleague of mine, Herbert Barry III, has analyzed cross-cultural data and found that the amount of crying reported for children seems to be related to the way in which the society reacts to their tears (35).

I have emphasized the possible role of learning in producing the observed individual differences in visceral responses to stress, which in extreme cases may result in one type of psychosomatic symptom in one person and a different type in another. Such learning does not, of course, exclude innate individual differences in the susceptibility of different organs. In fact, given social conditions under which any form of illness will be rewarded, the symptoms of the most susceptible organ will be the most likely ones to be learned. Furthermore, some types of stress may be so strong that the innate reactions to them produce damage without any learning. My colleagues and I are currently investigating the psychological variables involved in such types of stress (36).

THERAPEUTIC TRAINING

The experimental work on animals has developed a powerful technique for using instrumental learning to modify glandular and visceral responses. The improved training technique consists of moment-to-moment recording of the visceral function and immediate reward, at first, of very

small changes in the desired direction and then of progressively larger ones. The success of this technique suggests that it should be able to produce therapeutic changes. If the patient who is highly motivated to get rid of a symptom understands that a signal, such as a tone, indicates a change in the desired direction, that tone could serve as a powerful reward. Instruction to try to turn the tone on as often as possible and praise for success should increase the reward. As patients find that they can secure some control of the symptom, their motivation should be strengthened. Such a procedure should be well worth trying on any symptom, functional or organic, that is under neural control, that can be continuously monitored by modern instrumentation, and for which a given direction of change is clearly indicated medically—for example, cardiac arrhythmias, spastic colitis, asthma, and those cases of high blood pressure that are not essential compensation for kidney damage (37). The obvious cases to begin with are those in which drugs are ineffective or contraindicated. In the light of the fact that our animals learned so much better when under the influence of curare and transferred their training so well to the normal, nondrugged state, it should be worth while to try to use hypnotic suggestion to achieve similar results by enhancing the reward effect of the signal indicating a change in the desired direction, by producing relaxation and regular breathing, and by removing interference from skeletal responses and distraction by irrelevant cues.

Engel and Melmon (38) have reported encouraging results in the use of instrumental training to treat cardiac arrhythmias of organic origin. Randt, Korein, Carmona, and I have

had some success in using the method described above to train epileptic patients in the laboratory to suppress, in one way or another, the abnormal paroxysmal spikes in their electroencephalogram. My colleagues and I are hoping to try learning therapy for other symptoms—for example, the rewarding of high-voltage electroencephalograms as a treatment for insomnia. While it is far too early to promise any cures, it certainly will be worth while to investigate thoroughly the therapeutic possibilities of improved instrumental training techniques.

References and Notes

1. *The Dialogues of Plato*, B. Jowett, Transl. (Univ. of Oxford Press, London, ed. 2, 1875), vol. 3, "Timaeus."
2. BICHAT, X., *Recherches Physiologiques sur la Vie et la Mort* (Brosson, Gabon, Paris, 1800).
3. CANNON, W. B., *The Wisdom of the Body* (Norton, New York, 1932).
4. ALEXANDER, F., *Psychosomatic Medicine: Its Principles and Applications* (Norton, New York, 1950), pp. 40-41.
5. KIMBLE, G. A., *Hilgard and Marquis' Conditioning and Learning* (Appleton-Century-Crofts, New York, ed. 2, 1961), p. 100.
6. SKINNER, B. F., *The Behavior of Organisms* (Appleton-Century, New York, 1938); MOWRER, O. H., *Harvard Educ. Rev.* 17, 102 (1947).
7. MILLER, N. E. and DOLLARD, J., *Social Learning and Imitation* (Yale Univ. Press, New Haven, 1941); DOLLARD, J. and MILLER, N. E., *Personality and Psychotherapy* (McGraw-Hill, New York, 1950); MILLER, N. E., *Psychol. Rev.* 58, 375 (1951).
8. MILLER, N. E., *Ann. N.Y. Acad. Sci.*

- 92, 830 (1961); ———, in *Nebraska Symposium on Motivation*, M. R. Jones, Ed. (Univ. of Nebraska Press, Lincoln, 1963); ———, in *Proc. 3rd World Congr. Psychiat., Montreal, 1961* (1963), vol. 3, p. 213.
9. ———, in "Proceedings, 18th International Congress of Psychology, Moscow, 1966," in press.
10. ——— and CARMONA, A., *J. Comp. Physiol. Psychol.* 63, 1 (1967).
11. TROWILL, J. A., *ibid.*, p. 7.
12. MILLER, N. E. and DiCARA, L. V., *ibid.*, p. 12.
13. DiCARA, L. V. and MILLER, N. E., *Commun. Behav. Biol.* 2, 19 (1968).
14. ———, *J. Comp. Physiol. Psychol.* 65, 8 (1968).
15. ———, *ibid.*, in press.
16. EHRLICH, D. J. and MALMO, R. B., *Neuropsychologia* 5, 219 (1967).
17. "It even becomes difficult to postulate enough different thoughts each arousing a different emotion, each of which in turn innately elicits a specific visceral response. And if one assumes a more direct specific connection between different thoughts and different visceral responses, the notion becomes indistinguishable from the ideomotor hypothesis of the voluntary movement of skeletal muscles." [W. James, *Principles of Psychology* (Dover, New York, new ed., 1950), vol. 2, chap. 26].
18. MILLER, N. E. and BANUAZIZI, A., *J. Comp. Physiol. Psychol.* 65, 1 (1968).
19. BANUAZIZI, A., thesis, Yale University (1968).
20. MILLER, N. E. and DiCARA, L. V., *Amer. J. Physiol.* 215, 677 (1968).
21. CARMONA, A., MILLER, N. E., DEMIERRE, T., in preparation.
22. CARMONA, A., in preparation.
23. DiCARA, L. V. and MILLER, N. E., *Commun. Behav. Biol.* 1, 209 (1968).
24. ———, *Science* 159, 1485 (1968).
25. ———, *Psychosom. Med.* 30, 489 (1968).
26. DiCARA, L. V. and WEISS, J. M., *J. Comp. Physiol. Psychol.*, in press.
27. DiCARA, L. V. and MILLER, N. E., *Physiol. Behav.*, in press.
28. MILLER, N. E., *Science* 152, 676 (1966).
29. CARMONA, A., thesis, Yale University (1967).
30. For somewhat similar work on the single-cell level, see J. Olds and M. E. Olds, in *Brain Mechanisms and Learning*, J. Delafresnaye, A. Fessard, J. Konorski, Eds. (Blackwell, London, 1961).
31. KATKIN, E. S. and MURRAY, N. E., *Psychol. Bull.* 70, 52 (1968); for a reply to their criticisms, see A. Crider, G. Schwartz, S. Shnidman, *ibid.*, in press.
32. MILLER, N. E., DiCARA, L. V., WOLF, G., *Amer. J. Physiol.* 215, 684 (1968).
33. MILLER, N. E., in *Personality Change*, D. Byrne and P. Worchel, Eds. (Wiley, New York, 1964), p. 149.
34. LACEY, J. I. and LACEY, B. C., *Amer. J. Psychol.* 71, 50 (1958); *Ann. N.Y. Acad. Sci.* 98, 1257 (1962).
35. BARRY III, H., personal communication.
36. MILLER, N. E., *Proc. N.Y. Acad. Sci.*, in press.
37. Objective recording of such symptoms might be useful also in monitoring the effects of quite different types of psychotherapy.
38. ENGEL, B. T. and MELMON, K. T., personal communication.
39. The work described is supported by U.S. Public Health Service grant MH 13189.

INTRODUCTION

People are very good at discovering what is expected of them and trying to please by living up to such expectations. This simple fact of life is a recurring source of error to the behavioral scientist. If the experimenter has an idea about what the outcome of an experiment will be—and experimenters usually do—then the odds are very good that the subjects he interacts with will somehow sense the experimenter's expectations and intentionally or unintentionally change their behavior so that they "do what is right." The following selection by Robert Rosenthal was originally written as an introduction to the book *Clever Hans*. That book describes the behavior of a horse who, under very close scrutiny, gave every appearance of being an intellectual marvel. After prolonged investigation it became clear that the horse's abilities were indeed most remarkable but very limited; he was able to sense when the questioner—any questioner—expected him to stop tapping his foot. Since the questioners expected Hans to stop at the correct time—was he not the cleverest horse in history?—Hans sensed small movements which they made and *did* stop at the correct time.

As you read this article think how pervasive the problem is. Whenever we are trying to find out about somebody else, he, at the same time, whether intentionally or not, is probably changing his behavior so as to more closely match our expectations. There is no doubt that an understanding of such effects is going to be absolutely necessary if we are to improve our comprehension of human behavior.

CLEVER HANS: A CASE STUDY OF SCIENTIFIC METHOD

Robert Rosenthal

I

Ever since the Byzantine Empire was ruled by Justinian (A.D. 483–565) there have been reports of learned animals. But no animal intelligence so captured the imagination of layman and scholar alike as that attributed to Clever Hans, the horse of Mr. von Osten. Hans gave every evidence of being able to add and subtract, multiply and divide—operations performed with equal accuracy upon integers or fractions. He was also able to read and spell, to identify musical tones, and to state the relationship of tones to one another. His preferred mode of communication with his questioners was by means of converting all answers into a number and tapping out these numbers with his foot.

What Clever Hans and Mr. von Osten did for science (and for love), Clever Rosa, the mare of Berlin, and her owner did for entertainment (and for money). Any expert on animal training could observe that Rosa's

behavior was under the strict control of her trainer. Rosa gave correct answers because her trainer had only to bend forward in order to stop her tapping at the correct point, but no such signals could be observed to control Hans's tapping responses. In fact, on September 12, 1904, thirteen men risked their professional reputations by certifying that Hans was receiving no intentional cues from his owner or from any other questioner. Furthermore, these men, including in their number a psychologist, a physiologist, a veterinarian, a director of the Berlin Zoo, and a circus manager, certified that their investigation revealed no presence of signs or cues of even an unintentional nature. "This is a case," the investigating committee wrote, "which appears in principle to differ from any hitherto discovered." A "serious and incisive" inquiry into the cleverness of Hans was recommended and subsequently conducted. This book is the report by Oskar Pfungst of the procedure and the results of that inquiry which was undertaken in collaboration with Carl Stumpf, the eminent psychologist.¹

Clever Hans constitutes a famous "case" in the history of psychology. One wishes to know about it for its own sake. There is, however, another reason for reprinting this book, long

Reprinted with permission of the author and publisher from "Introduction" by Robert Rosenthal to Clever Hans by Oskar Pfungst, edited by Robert Rosenthal. Introduction copyright © 1965 by Holt, Rinehart & Winston, Inc. Reprinted by permission of Holt, Rinehart & Winston, Inc.

so difficult to obtain. The investigation of Hans's abilities is also a classic in respect of method, for it is a first-rate scientific inquiry into what began as an incomprehensible case history.

Pfungst's work is ideal as a paradigm for this kind of inquiry. First, he did careful and sophisticated work. Second, he described his procedures and his thinking in such detail that anyone, layman or student, can understand what goes on in a scientific inquiry. Third, the subject of his investigation has great intrinsic interest.²

First Pfungst established that Hans was, in fact, clever, and that his cleverness did not depend on the presence of his master, Mr. von Osten. Anyone, or almost anyone, could put a question to Hans, and the chances were good that an accurate answer would be forthcoming. Next, Pfungst employed a control condition in fitting Hans with blinders so that he could not see his questioners. This experimental addition reduced Hans's cleverness. There was another experimental change that cramped Hans's style: The questioners asked questions to which they did not know the answers. Pfungst also noted that Hans's accuracy diminished as the physical distance between him and his questioners increased. The closer the questioner, the more impressive was Hans's performance. From these data Pfungst could see that Hans was clever only when he had visual access to a source of the correct answer. In this indirect manner the likelihood had been increased that visual cues were involved, despite the fact that no specific visual cues had as yet been discovered.³

Eventually Pfungst noticed that the slight forward inclination of the

questioner's head, to better see Hans's hoof tapping, was the signal for Hans to begin tapping. This forward inclination of the head did not need to be accompanied by a question for it to serve as a stimulus to tapping by Hans. When a correct answer required a long series of taps Hans would tap at a faster rate, as though he knew he had a lot of work ahead of him. This accommodation to the task at hand, which greatly added to the intellectual status ascribed to Hans, also came to be understood by Pfungst: He found that the interrogator, when putting a question to Hans, tended to lean further forward when the number that was the answer was a large one—as though he were settling down for a long wait. Hans's rate of tapping apparently depended on the angle of the questioner's forward inclination. Similarly, the cue for Hans to stop tapping also turned out to be derived from the questioner's expectancy for a correct answer; for when Hans reached the correct number of taps, the questioner tended to straighten up—and that was the cue for Hans to stop. Hans was sensitive to tiny upward motions of the head, even to the raising of eyebrows or the dilation of nostrils, any of which was sufficient to stop his tapping. Pfungst showed that anyone could start Hans tapping, and that anyone else could then stop his tapping at any time by the use of these cues. Most interesting was the finding that even after he had learned the cueing system very well Pfungst still cued Hans unintentionally, though he was consciously trying to suppress sending the crucial visual messages.

When Pfungst had learned all this about this man-horse communication system, he did a very modern thing indeed. He took his discoveries to the

laboratory for further confirmation. As he put it, "Thus, artificial synthesis became the test of the correctness of analytical observation." In order to increase the generality of his findings at this point he not only increased his sample size but sampled from a different species—people, instead of horses. Pfungst took the part of Hans, inviting questions to which he tapped out his answers.

Of 25 subjects, 23 unintentionally cued Pfungst as to when he should stop tapping! When errors did occur, they tended to be errors of either one too many or one too few taps—the kind of errors Hans was most likely to make. The subjects, including men and women of assorted ages and occupations, had of course not been told the intent of the experiment. One of the subjects was a psychologist trained in introspection, but even he was unable to discover that he was emitting cues to Pfungst, Hans's surrogate.

To record his subjects' responses permanently and for leisurely study, as well as to amplify their very subtle cues, Pfungst developed an ingenious system of instrumentation. The questioner-subjects were hooked into a series of recording devices which would record an amplification of their head movements on three planes. In addition he undertook to measure the changes in his subjects' respiration. In interpreting the results of these recordings, Pfungst postulated a build-up of tension in his subjects that was released as a motor response to the perception of the correct answer as the experimenter tapped it out.

II

Pfungst's findings solved not only the riddle of Clever Hans but in prin-

ciple the problem of other "clever" animals. Now there was a reasonable explanation available for the ability of Hugo Kretschmer's dog to answer questions by means of pressing a bell. Similarly, the abilities of the English bulldog "Kepler" of Sir William Higgins could be better understood. Kepler barked the answers to problems put to him, just as Hans had tapped them out. It seemed only reasonable to assume that by some tone or movement Sir William started Kepler's clever barking, and then by some other subtle movement stopped his dog's barking at the number of barks that were correct. Most of the other clever animals of that time were never so mysterious because their cueing was obvious, and often obviously intentional. That was the case with Hans's vaudeville rival, Rosa, the mare of Berlin. Similar mechanisms seemed to operate in the clever dog of Utrecht, the reading pig of London, and the reading pig and talking horse of Virginia.^{4, 5}

In his Preface to the present volume, James R. Angell mentioned its relevance to the then current wave of interest in animal life and behavior—a relevance that is, if anything, even greater today. Psychologists, behavioral scientists, and ethologists are more than ever interested in animal behavior as it occurs in the field as well as in the laboratory, and "a dominant theme in the remarkable growth of interest in animal behavior during the past decade . . . has been the study of animal communicative systems . . ." (Alexander, 1964, p. 713). Not horse, dog, and pig so much now, perhaps, but insect and octopus, gorilla and rhesus, dolphin and mynah bird are objects of intensive study. The layman, too, shares in the increased interest in animal behavior

through the publication of books that are at once scholarly and popular (for example, Broadhurst, 1963). Broadhurst mentions the case of Clever Hans as illustrating Lloyd Morgan's canon—the principle that we should infer in animals only the lowest level of mental ability that is consistent with their observed behavior.

Angell believed that Pfungst's work was also relevant to the interest then invested in "telepathy" and "muscle reading." Only a few years after Pfungst's investigation of Clever Hans's ability, the American psychologist Stratton (1921), who had been a student of Wundt's, undertook a similar investigation of a famous muscle reader of the time.⁶ Eugen de Rubini was able to read the "thoughts" of others by reading their patterns of muscle tension. All the mystery and excitement surrounding the horse of Mr. von Osten emerged for de Rubini's amazing abilities. And as Hans had his Pfungst to make his case one of scientific importance, so did de Rubini have his Stratton. Stratton's investigation was similar in strategy to that of Pfungst. The psychologist E. C. Tolman was present throughout all the experiments, and another psychologist, Warner Brown, was present during most of them.⁷ The subject, de Rubini, held one end of a watch chain and his guide held the other end. The guide was to think of some object in the room that would then be identified by de Rubini. Success, which occurred much more often than it should by chance, was the more remarkable because the watch chain's slackness would seem to reduce the likelihood of muscle tension being transmitted. Although de Rubini was a professional performer he made no claims of supernatural power and was quite willing to be tested under

controlled conditions after the manner of Mr. von Osten. And, as in the case of Clever Hans, no observer was able to detect a gross sensory cue to de Rubini. As a matter of fact, de Rubini's performance tended to worsen if a guide intentionally tried to help him by giving him gross cues. When the muscle reader held the end of the slack watch chain his success rate was twice as high as would be predicted on the basis of chance alone. When he did not hold the watch chain his success rate dropped to about one and one-half times as compared with what might occur by chance. When auditory cues to the subject were reduced his performance actually increased, a fact suggesting that some background sounds interfered with his performance. To achieve maximal accuracy, de Rubini required only peripheral visual contact with his questioner-guide. Total elimination of visual cues reduced his successes to about the level of chance, but the nature of his errors was not random. He was closer to naming the right object more often than could be expected if he were obtaining no information at all from his guide. Clear failure of de Rubini's abilities occurred only when both visual and auditory cues were drastically reduced. Thus we can see that Pfungst's strategy for investigating the case of Clever Hans was largely taken over by Stratton and his colleagues, Tolman and Brown, in their investigation of de Rubini. Unlike Pfungst, however, these investigators never did identify the precise cues employed by their subject.

Two years after Stratton's study of the muscle reader, W. S. Foster (1923) undertook a strategically similar series of experiments to investigate the role of unintentional communication in

the ability to divine the presence of water and metals. Just as Stratton had been assisted by Tolman and Brown, so was Foster assisted by two eminent psychologists, R. M. Elliott and D. G. Paterson.

The subject of their study was a 70-year-old pastor of a church located in a small town near the University of Minnesota. He had a 45-year history of successfully locating not only water, gold, and silver but oil, natural gas, and iron as well. His feats had been observed by respectable and tough-minded critics, including professional engineers. That there was nothing fraudulent about his performance had been inferred from his vocation and from the fact that he requested the Psychological Laboratories at the University of Minnesota to study his rod-divining powers scientifically. He shared with von Osten and de Rubini a willingness to be "exposed" if that were to be the verdict of the scientists.

One of the experiments undertaken in order to stimulate field conditions required the subject to locate the course of water mains leading to two buildings on the campus. The rod diviner was accompanied by an experimenter who knew the location of the water mains in question. The results of this study suggested that the subject was not only able to locate the path of the water mains more accurately than could reasonably be attributed to chance but that he could also gauge the depth of the pipes. Foster's interpretation of this result was that quite without intention the old man was able to read subtle cues unintentionally emitted by the accompanying experimenter.

For another experiment a small cardboard box was filled with silver coins and silver and gold watches.

The task in this experiment was for the subject to divine the depth of these materials. The pastor stood in a second floor hallway holding a 20-inch wire divining rod. In the first-floor hallway, immediately beneath the pastor, a 13-rung ladder was placed. By varying the rungs on which the materials to be located were placed, the depth of the metals relative to the subject could be varied at random. On a landing between the first and second floors a group of observers was placed in such a way that they could see both the ladder on the first floor and the rod diviner on the second floor. In spite of the fact that the spectators had been cautioned to avoid emitting cues to the subject, and in spite of the subject's sincere inattention to the spectators, he was able to gauge the depth of the metals far better than chance would allow. In fact, the correlation between the actual and the divined depth was found to be $+0.77$ for a series of 26 trials. It seemed likely that the unintended and involuntary movements by the spectators served the necessary cueing function in this series of trials. Another possible source of cues to the subject in this study was a faintly audible sound from the first-floor hallway associated with the experimenter's placing and replacing materials on the various rungs of the ladder. These sounds were, however, probably even less audible when masked by the slight noise associated with the group of spectators.

When the same task was set for the rod diviner *without* any observers present who knew the correct answer the subject's performance fell off dramatically, although it did not seem to disappear. The correlation between the actual and divined depth fell from the previous $+0.77$ to $+0.29$. This

somewhat better-than-chance accuracy seemed best interpreted as due to the subject's hearing the faint noise of the experimenter below climbing up and down the ladder.

As a more definitive test of the rod diviner's accuracy the following experiment was carried out. On a large and sturdy table 14 squares were marked off in alignment with 14 squares on the floor below. Cardboard screens shielded the squares under the table from observation. An experimenter placed the box of coins and watches in one of the randomly selected squares beneath the table and then left the room. Each square was selected twice for a total of 28 trials. The experimenter who recorded the subject's divinations did not know the location of the hidden materials. Under these conditions the pastor, standing on the table, was totally unable to locate the coins and watches. As in the case of Clever Hans, when no one was about who knew the correct answer and could unintentionally clue the subject, performance fell to a chance level.

When Foster instituted another table-top experiment in which the amount of material was varied, and in which the experimenter who knew the correct amount was present, performance again began to improve to a nearly significantly better-than-chance level. In this series of trials the experimenter said only the word "ready" to the subject, after having varied the amount of material in the cardboard box. Foster felt that the tone in which this single word was uttered might have unintentionally communicated some information about the right answer to the subject.

Involuntary cueing of subjects in experiments in extrasensory perception (ESP) has long been discussed. In

1898, Moll repeated Alexander Wernicke's still earlier warning that muscle tremors on the part of the sender might influence the response of the receiver. Some 40 years after this and similar warnings had been sounded Kennedy (1938, 1939), in an improved replication of a much earlier experiment (Hansen and Lehmann, 1895), was able to show by the use of parabolic sound reflectors that senders in experiments on telepathy could unconsciously whisper or otherwise signal by auditory cues the right answers to the receivers. All this occurs, of course, not only in complete innocence after the manner of Mr. von Osten but also of Pfungst, Stumpf, Stratton, Holman, Brown, Foster, Elliott, and Paterson. The subtlety of cues sufficient to inform subjects of the correct response was amply demonstrated by J. G. Miller (1942), who showed that subjects could perceive subtle cues at levels of illumination ordinarily found to be subthreshold.

Angell's comment in his Preface to this volume that Pfungst's findings would have relevance to an understanding of ESP and muscle reading was not only accurate but also, as it proved, prophetic.

"Talking" and "reading" horses, dogs, and pigs; clairvoyants, muscle readers, and rod diviners—the very mention of these oddities is enough to make the blood boil in the veins of some of the tough-minded psychologists of today. The phenomena are still too imperfectly analyzed to receive important consideration in the modern course in introductory psychology. Nevertheless, these topics have engaged some of the best minds of the day, and sophisticated research techniques were employed by these men to study their phenomena. The classical conditioning of Pavlov seems

a more solid field; it was Pavlov who noted the unintentional communication process between laboratory experimenter and laboratory animal (Razran, 1959; Zirkle, 1958). Pavlov had been interested in the Lamarckian problem of the inheritance of acquired characteristics and had conducted some research in this area. It seemed that successive generations of trained mice were better able to learn, but at the Thirteenth International Physiological Congress held in Boston in August, 1929, Pavlov offered a fascinating postscript to his puzzling data. He reported that a closer check of his data suggested that there was no such increased learning ability on the part of the mice; instead there was an increased teaching ability on the part of his research assistant (Gruenberg, 1929). Unfortunately, from this interesting information alone we are in no position to state just how the experimenter unintentionally instructed his mouse subjects.

There was also the important work of H. M. Johnson (1913) on auditory discrimination and learning in dogs.⁸ He presented an extended critique of Kalischer's earlier work, which indicated that his dogs could correctly anticipate the experimenter's next response from his posture, respiration, and the pattern of strain and relaxation of the muscles of the head and body. These cues too were obviously of an unintended and involuntary nature, as had been the case with Hans's questioners. And, as in their case, the experimenter's expectations were likely to be the antecedent conditions of an unintended communication process. Johnson also cited the work of Max Rothmann, who, as early as 1908, only one year after the appearance of Pfungst's work with Clever Hans, believed Kalischer's dogs

to have reacted in the same manner as Hans. As a control for the Clever Hans phenomenon, Rothmann employed as an experimenter a servant who was unaware of the purpose of the investigation. That even under these conditions dogs could profit from the experimenter's unintended cues did not surprise Johnson. Dogs could, after all, learn a servant's signaling system as easily as a psychologist's. Johnson himself introduced the necessary controls into this situation by using a series of experiments in which the dogs could not see him at all, presumably eliminating visual cues effectively. To partially control auditory cues Johnson suggested that the experimenter not watch the dog's responses. Thus the behavior of the dogs could not elicit a change in the vocal behavior of the investigator, whose expectations might have led to changes in his respiration as a result of his observation of the dogs' responses. When all the proper controls had been instituted, Johnson found that no discrimination abilities at all could be demonstrated for the dogs.

That minimal cues, unintentionally emitted by the experimenter, might influence the subject's response has also been suggested for psychophysical experiments. In an experiment on the judgment of weights, Warner and Raible (1937) removed the experimenter from the visual field of the subject. In half the trials the experimenter knew which answer was correct, and in half he was kept uninformed. Only 6 of the 17 subjects showed a large discrepancy from a chance distribution of errors, but all of the 6 subjects made more errors on those trials in which the experimenter did not know the correct answers himself. The authors suggested that some form of auditory cue might have been

emitted by the experimenter. That this cue was not likely to be vocal was suggested by the fact that the experimenter kept his mouth tightly closed in order to prevent unconscious whispering or any other unintended vocal signaling. These investigations, like most of the others just discussed, did not succeed in isolating (as Pfungst had been able to do) the specific cues employed.

A paradigm similar to that used by Warner and Raible, and one very much in the tradition of Clever Hans, was chosen by Stanton and Baker (1942). In a study of the retention of geometric figures 200 undergraduate students were tested by five experienced workers. These experimenters had with them a key of the "correct" responses, some of which were intentionally listed incorrectly. Experimenters were explicitly cautioned to avoid any biasing effect of the knowledge of the "correct" responses. Nevertheless, results of the retention test showed that, when the incorrect answer had been substituted for the correct one on the experimenter's key, his subjects were more often "correctly" incorrect in their response. By some cue or other, experimenters must have communicated their expectation to their subjects. In spite of the fact that Lindzey (1951) and Friedman (1942) were unable to replicate these findings, the results were so striking and so consistent with some of the earlier findings described here that it seems improper to dismiss them. In fact Stanton (1942a) later presented additional evidence to strengthen his conclusion. Again employing nonsense materials, he had three groups of experimenters test **subjects for retention**. One group of **experimenters knew** which of the materials had been shown to subjects

earlier. Another group of experimenters was intentionally misinformed as to which materials had been shown to subjects, and a third group was told nothing. Results of this experiment also indicated that experimenters had in some way communicated their knowledge of the correct answers to their subjects—subtly and unintentionally.

A final example of the unintentional cueing of subjects is the case briefly reported by Wilson (1952). The experimental task required an apparently difficult discrimination between the presence and absence of a faint light. Ability to discriminate appeared to depend in part on the experimenter's recording system. It turned out that maximal discrimination occurred when the experimenter, using a scratchy pen, unintentionally cued the subject by a short pen scratch for no light and a long pen scratch for the light.

It is interesting to compare the results of the investigations described with the results of Pfungst's experiments. All these investigators learned something about the experimenters' unintentional communication with their subjects. But with the possible exception of the case described by Wilson (1952)—which seems to have been a much simpler one than any of the others—no one but Pfungst succeeded in specifying the exact cues that had been unwittingly emitted to the subject. Johnson had been able to implicate the experimenter's respiratory pattern, but he could not be sure of it. Stratton, Foster, and Warner and Raible were able to state only the sense modalities likely to be involved in the unintended communication. Pavlov, and Stanton and Baker, could not specify the modalities involved at all. Of all the com-

plicated cases of unintentional communication studied, only Pfungst's case can be regarded as having been completely solved. Taking all the cases together does suggest, however, that ordinary, though often extremely subtle, sensory cues may be sufficient to explain the unwitting communication between experimenters and subjects. Some more recent attempts, both to elicit and to elucidate such unwitting communication, will be described in the following section.

III

In many of the cases of the Clever Hans phenomenon the subtle and unintentional cueing of the subject by the experimenter could be attributed to the experimenter's expectancy. Pfungst, for example, considered expectation and its attendant tension as helping him to account for the communication of correct responses to Clever Hans (see p. 147).

In general Hans performed accurately only for those questioners who consciously or unconsciously believed he could. Merton's (1948) concept of the self-fulfilling prophecy seems relevant: one behaves in such a way as to increase the likelihood of the prophesied event. That many experimenters over the years may have fulfilled their experimental prophecies by unintentionally communicating information to their subjects may be a disquieting proposition. We will probably never know what proportion of behavioral researches have been affected by the unintended communication of experimenter expectancies. This problem, the Clever Hans phenomenon, has occupied the present writer's attention for most of the last decade.

With horses so hard to come by and

troublesome as laboratory animals, rats will do. A dozen students of experimental psychology served as experimenters in a study of the learning of a brightness discrimination in which rats were purported to have been specially bred for cleverness (Rosenthal and Fode, 1963). Each experimenter trained five animals a day for five days. Half the experimenters were told that their animals were from the "Berkeley Colony: Maze Bright Division," and the remaining experimenters were told that their animals were from the "Maze Dull Division." The experimenters' expectations were thus manipulated by alleged reference to the subjects' genetic history. All the rats were, of course, from a homogeneous population and were assigned at random to experimenters. At the end of the first day's running, the animals used by experimenters expecting better performance outperformed the animals used by experimenters expecting poorer performance. The difference found was statistically significant and held up over the full five-day schedule. Animals believed to be bright made half again as many correct responses as those believed to be dull. The particular discrimination employed was a difficult one, and the animals often refused to make any response at all. Moreover, among animals used by experimenters expecting good performance no running at all occurred on 11% of the trials, whereas the animals used by experimenters expecting poor performance refused to run on 29% of the trials.

What cues could have been given the rats by their experimenters to account for the results of this experiment? It would be preposterous to suppose that the rats carefully observed the head movements of the ex-

perimenters as Hans had done, or listened for auditory cues as John-son's dogs may have done. Still cues can operate, as it were, automatically. To this writer it seems more likely that these experimenters handled their rats differentially as a function of their expectancy about their abilities and that the differences in handling these animals led to the differences in their performance. How does one handle a rat believed to be bright? Perhaps more often and more gently; perhaps more "warmly" and enthusiastically than one handles an animal believed to be dull. Associated with the perception of animals' brightness in our sample of experimenters was the perception of animals' "character." Rats believed to be brighter were described by their experimenters as cleaner, tamer, and generally more pleasant. It may have been these associated characteristics that led the experimenters to handle their brighter-perceived animals differentially. In any case, experimenters who believed their subjects to be bright stated that they did handle these subjects more often and more gently than did the experimenters who expected less of their subjects.⁹

If, as Pfungst and others found, the expectancies of experimenters are communicated so readily to their animal subjects, we might expect that they can also be communicated to their human subjects. For this reason a series of studies was undertaken to learn the role of the experimenters' expectancy in influencing the responses of their human subjects. A task of person-perception was employed in many of these investigations (Rosenthal, 1963). Subjects were asked to judge what degree of success had been experienced by each person pictured in a series of photographs. The

subjects rated each photo from very unsuccessful (-10) to mildly unsuccessful (-1) to mildly successful (+1) to very successful (+10). The photos, cut from a magazine, had been standardized to evoke average ratings of zero (neither successful nor unsuccessful). The basic procedure employed in many of these studies was to draw a random sample of experimenters and to lead half of them to expect that the subjects assigned to them would tend to see the people pictured as failures. The remaining experimenters were led to expect that their subjects would tend to see the people pictured as successes. Actually, of course, subjects were simply assigned to their experimenters at random. All experimenters were given identical instructions to read to their subjects. Nevertheless, in three studies in which 30 experimenters each used about a dozen subjects, the lowest average rating obtained by any experimenter expecting high ratings was higher than the highest average rating obtained by any experimenter expecting low ratings from his subjects. In all three studies these nonoverlapping distributions of experimental results differed very significantly from any results that could be attributed to chance.

The conclusion to be drawn from these and subsequent studies (Rosenthal, 1964) is that in some subtle, unintentional way experimenters do communicate their expectations to their human subjects, whose performance is then significantly altered. There remains, however, the most urgent question: How do the experimenters communicate to their subjects what it is they expect of them?

The exquisite sensitivity of human subjects to these "demand characteristics" has been amply documented

by Martin Orne (1962). Data from several experiments, including a careful analysis of sound motion pictures of interactions between the experimenter and his subject, suggest that no gross procedural errors are responsible; neither words nor readily perceptible gestures carry the burden of communication. The explanation does not lie in the infrequent occurrence of biased errors of observation, recording, and computation.

Now it seems reasonable to suppose that the cues employed by experimenters who intentionally tried to influence their subjects might appear as simple exaggerations of more subtle cues employed when the experimenters were influencing their subjects unintentionally. Accordingly, sound motion picture films were made of three experimenter-subject interactions in which the photo-rating task was employed. Each experimenter was assigned at random a number between -10 and +10. It was made the experimenter's task to influence his subject to rate the stimulus photos as close to the randomly assigned numerical value as possible. The three films were then viewed by a large group of observers who, of course, were not told what rating each experimenter was trying to elicit from his subject. All observers made an effort to guess the numerical value each of the three experimenters was trying to obtain from his subject. These guesses proved to be so accurate there can be little doubt that the observers accurately "read" an experimenter's expectancy when he intentionally tried to influence his subject's responses.¹⁰

When observers can agree so well on an experimenter's expectancy, we would expect them also to agree on the channel of communication in-

involved. Numerous hypotheses were suggested by the observers, but there were few agreements on the specific cues employed. There appeared, however, to be two promising dimensions for the location of hypotheses. One of these had to do with timing; that is, when in the social interaction the communication occurred. The other had to do with sense modality; that is, whether the cues were more visual-kinesic or more auditory-paralinguistic.¹¹ For many observers the communication of expectancy was judged to occur only after the subject began making his responses. These observers were our "reinforcement theorists," who pointed out that when subjects gave a response close to the desired value, the experimenter responded with positive reinforcing cues such as smiling, head-nodding, looking more pleased or interested, and recording the responses more vigorously. The reinforcement theorists pointed out, moreover, that when subjects gave responses not in accordance with the expectancy, the experimenters employed negative reinforcing cues—head-shaking, eyebrow-raising, looking surprised or disappointed, looking the response aloud, tapping the pencil, holding the stimulus photo up longer, tilting the photo forward, and "throwing" the photo down on the table. There was good agreement among observers that no two experimenters showed exactly the same pattern of reinforcements.

Opposed to the reinforcement theorists were observers who felt that communication of expectancy preceded subjects' response to the experimental task. These theorists were divided as to the nature of the cues involved. One type, the specifist, emphasized the potential cues associated with the experimenter's reading of the

instructions to the subject. These observers believed they had noted a greater emphasis on the desired alternative when the subject's alternative responses were described for him by the experimenter, but there was little agreement on the form of this emphasis. Vocal stress in speech was suggested as one form of emphasis, but so were stammering, changing the rate of speech, making more reading errors, and pointing a little longer at that region of the rating scale from which the subject was to select his response.

The other type of nonreinforcement theorist was the generalist, who tended to de-emphasize the importance of any single, small cues. Instead these observers emphasized the general climate created by the experimenter from the very beginning of his interaction with his subject. Thus, the experimenters who expected more positive ratings from their subjects were seen as creating a more positive tone, whereas those who expected negative ratings were seen as creating a more negative tone. As part of this general tone the experimenters' ocular dynamics were specifically noted: greater eye-contact was associated with positive tone and visual-avoidance with negative tone.

The great differences in what the observers saw on looking at the intentional communication of the experimenter's expectancy are reminiscent of the early search for the key to communication of information to Clever Hans. In this case, however, the analysis is not as yet at hand—neither for intentional nor for unintentional communication.

In his Introduction to the present volume Stumpf confessed that he had expected to find auditory cues for Hans rather than minute visual cues

—cues so subtle that even expert observers who expected them could not detect them. Still Stumpf was not altogether wrong, for in at least some tasks it appeared that auditory cues were of some value to Hans. When Hans was asked to fetch placards with certain words written upon them or to select one of several colored cloths, Mr. von Osten would shout at him if he approached the wrong stimulus—thereby directing him to the correct one. In experimental observation of this task Pfungst found Hans to be about 30% accurate with benefit of visual cues alone. When auditory cues were added, Hans's performance jumped to over 50% accuracy.

Following closely in Pfungst's footsteps, Fode (1960) undertook a study directed at learning the relative contribution of visual and auditory cues to the process by which an experimenter unintentionally communicates his expectancy to his subjects. The standard photo-rating task was administered by 24 experimenters to 180 men and women subjects. The 24 experimenters were randomly assigned to one of four groups. One group of six experimenters, the *control* group, was led to expect their subjects to rate the photos as being of unsuccessful people. The remaining three experimental groups of six experimenters each were led to expect their subjects to rate the photos as being of successful people. The experimental groups differed from each other in the types of cues they unwittingly sent their subjects. The experimenters permitting *visual cues* were visible to their subjects, but remained entirely silent except for greeting their subjects and handing them written instructions. The experimenters permitting *auditory cues* read their instructions to subjects aloud, but were hidden from

their subjects' view by an interposed screen. Those free to give both *visual and auditory cues* read their instructions to their subjects and also remained in full view. This group differed from the control group only in the induced expectancy.

The results of this study showed no difference in obtained photo-ratings between the group permitted visual cues and the control group, a result which suggested, among other interpretations, that visual cues alone were insufficient to carry the burden of the unintended communication or that the strangeness to the subjects of an apparently unnecessarily aloof and mute experimenter led to their rating the photos as more unsuccessful. The group permitted auditory cues obtained significantly more ratings of success than did the control group, a result suggesting that auditory-paralinguistic cues alone might be sufficient to communicate experimenters' expectancies to subjects. The group permitted both visual and auditory cues obtained significantly more ratings of success than did the group permitted only auditory cues. It appeared, then, that the effects of auditory cues, although sufficient, could be significantly increased (approximately "doubled") by the addition of visual cues, even though the effect of visual cues alone was at best only equivocal. In any case, it seems clear that neither the strategy nor the tactics of inquiry employed by Pfungst are in any way outmoded or irrelevant to contemporary psychology.

Pfungst had noted among questioners of Hans that some were more successful than others in having their expectancies confirmed by Hans's responses. The configuration of the characteristics that Pfungst observed to be associated with more successful

unintentional influencers of Hans is very like that observed to be associated with more successful unintentional influencers of human subjects. Thus Pfungst discovered that an air of authority about the questioner increased the likelihood of his influencing Hans, and now current experiments show that an air of authority increases the effects of experimenters' expectancies upon their human subjects. Pfungst found that ability and tact in dealing with animals led to more expected responses from Hans. The current experiments show that an interpersonal style, reflecting a relaxed, personal, interested, and involved approach to human subjects, also tends to maximize the effects of an experimenter's expectancy. Pfungst found that experimenters who were more gesturally inclined were more able to influence Hans's responses. Similarly it now appears that experimenters more inclined to gesture are unintentionally more influential in determining their human subjects' responses.

Pfungst also noted even Hans's better questioners became more successful in eliciting the expected responses as they had more practice. A similar learning-effect has been observed in these studies of the experimenter's expectancy as it affects human subjects. For these, as for Hans's experimenters, their subjects' responding in accordance with their expectancy may well have served as reinforcement. Whatever small, unintended visual or auditory cues were produced by the experimenters would have then become more likely to recur when they increased the likelihood of obtaining an expected response. Pfungst's findings, as well as the present ones, may be interpreted to mean that not only does the experimenter unknowingly

evoke expected responses from his subject but that the subject also unwittingly evokes appropriate cues from his experimenter.

Pfungst and Stumpf summarized their difficulties in learning the nature of Clever Hans's talents by speaking of "looking for, in the horse, what should have been sought in the man." There may be many other areas of contemporary psychological inquiry in which what we have looked for in our *subjects* should have been sought in ourselves as *experimenters*.

IV

Are there methodological morals to be drawn from the case of Clever Hans? Whatever questions we may have to put to horse or man might best be asked by any collector of data who has no specific expectancy regarding what the answer ought to be. The ultimate collector of data, were we to insist on an absence of any expectation regarding subjects' responses, would have to be some sort of machine. Nor is this such a bizarre notion, for there exist now automated systems for the collection of data that drastically reduce—though they do not quite eliminate—contact between experimenter and subject. Someone must still place the animal into the automatic apparatus.

Considering that in much contemporary behavioral research experimenters have contact with at least two types of subjects, some from an "experimental" and others from a "control" group, the experimenter's knowledge of the subject's group membership is likely to affect his expectancy regarding their responses. In order that their expectancies should not affect the two kinds of subjects differentially, some investigators have taken

to recording their instructions to subjects on tape so as to assure standard presentation to subjects under different experimental conditions. That would seem to be a reasonable safeguard against the experimenter's tone of voice (and other paralinguistic features) influencing his subjects' responses differentially. We can well imagine that, if Pfungst had had tape recorders available, he would have recorded his questions for Hans, started the tape by remote control, simultaneously activating a motion picture camera directed at Hans's trained foot. If Hans had tapped at all, he would of course still be tapping with no visual cues to stop him. But he might not have tapped, and Mr. von Osten would then have complained that Hans's training had not included any audio-visual instruction and that to be realistic a human questioner must be employed—and Mr. von Osten would have been exactly right. Nor is this fantasy entirely irrelevant to contemporary psychological research. There are experiments waiting to be conducted, many of them in social psychology, which would simply make no sense at all without a present and fully operating human experimenter. In these cases it is still possible to assess the relative contribution to subjects' responses of the experimenter's expectancy even though it cannot be eliminated.

This result could be accomplished by the frequent employment of "expectancy control groups," and there are very few areas of behavioral research which would not benefit from such a technique. Pfungst utilized such expectancy controls when he employed some experimenters who, because they knew the right answer to the question put to Hans, expected him to stop tapping at a certain num-

ber, and other experimenters who had no such expectancy. We can generalize this procedure in the light of what we know both about the effects of the experimenter's expectancy and current methods of research.

Consider any experiment in which two different conditions are involved: one regarded as an experimental and the other as a control condition. Examples are a new teaching method vs. an old one; a new drug vs. a placebo; two types of psychotherapy; a salient and a nonsalient persuasive communication. The experiment is being conducted, in the first place, because we think one of these conditions will lead subjects to respond differently from the other. The subjects in one group will learn faster or lose their symptoms sooner or change their personalities more or alter their attitudes more. If we are concerned, as we should be, about the effects of expectancies on subjects' responses, we can introduce another experimental dimension in which we vary the expectancy of the data collector. Without necessarily increasing the total number of subjects involved we would merely divide the experiment into two halves. One half of the subjects in the "experimental" condition would be interviewed by data collectors who had been led to expect that the theoretical effect (change or improvement or difference) would occur. The other half of the subjects, though receiving the same programmed treatment (drug, communication, and so on) would, however, be seen by experimenters led to believe that there would be no effect—perhaps by a false labeling of the group or by disparagement of the effectiveness of the treatment. Similarly, among the subjects of the "control" condition, half would be interviewed by experimenters who

expected the effect of the experimental manipulation and half would be seen by experimenters who expected no effect. Appropriate statistical analysis of the results of such an expectancy-controlled experiment would permit an assessment of the magnitude of the effects of the experimental variables of primary interest in relation to the effects of the experimenter's expectancy.¹²

Appropriate to the discussion of methodological issues, but important in its own right, is Pfungst's interest in the various sense modalities by which man may communicate with animals. If we knew precisely by what means we unintentionally communicate our expectancies to our animal and human subjects, we could institute more effective controls against the effect of our expectancies. More generally, if we knew more about the modalities by which we subtly and unintentionally influence one another, we would then have learned a great deal that is new about human social behavior. Pfungst emphasized the visual and auditory sense modalities in his study of Clever Hans. Subsequent investigators have also tended to emphasize the role of these modalities in interpersonal influence. At present there appears in fact to be a renaissance of interest in the role of kinesic and paralinguistic information as it contributes to the total process of communication (Sebeok, Hayes, & Bateson, 1964).

Not so much in connection with Hans as with respect to horse and rider relations generally, Pfungst discussed the importance of kinesthetic and tactual cues. By means of these cues horses are virtually able to anticipate the wishes of the rider quite without the rider's awareness. The role of the skin senses in human com-

munication has recently been brought into focus by Geldard (1960), who has presented evidence for a remarkable sensitivity of the skin even to human speech. For some time it was supposed that the blind were enabled to avoid obstacles through an exquisite sensitivity of the skin of the face, called "facial vision," but it has been shown that the sense modality actually involved is auditory. Recent ingenious studies carried out by Kellogg (1962) suggest that the blind at least may employ a technique of echo ranging (sonar) to obtain information about their environment, much in the manner of the bat and the porpoise. Kellogg's human subjects were able to assess accurately the distance, size, and composition of objects in their environment.

One very neglected modality from the viewpoint of interpersonal communication is the olfactory. Vision and audition are useful in human communication, not only because these senses convey so much information about the environment, including the people in it, but also because these senses tell us much about even small changes in the interpersonal environment. Olfaction may or may not be capable of giving human beings this sort of information. Although little empirical research has been conducted on this question, it seems reasonable to assert that, because we are alert to such small changes in others' facial expressions and tones of voice, we could be sensitive to changes in odors of others.¹³ A great deal of research would seem to be needed here. When we know much more than we do now about communication by minimal cues of any sense modality, we will be in a far better position to understand the structure and dynamics of such interpersonal relationships

as those that occur in families, schools, businesses, clinics and hospitals—and in the collecting of data too.

Notes

1. The book which lies ahead in Pfungst's, yet in noting that Stumpf was both a collaborator of Pfungst and eminent as well, more should be said about him. What could and should be said has been said by Boring (1950). One of the most important (but clearly not *the* most important) of the German psychologists of the waning years of the last century, Stumpf nevertheless held the most important chair of psychology in all Germany, the one at the University of Berlin. Oskar Pfungst was one of his outstanding students. That Pfungst should have conducted the investigation of Clever Hans at Stumpf's suggestion made sense, for Stumpf often turned experimental enterprises over to students who would accept them. As Boring (1950) put it: "Stumpf was an experimentalist by philosophical conviction but not by temperament" (p. 371).

Stumpf, of course, was one of the thirteen courageous men who risked their professional reputations by absolving Hans from the reception of any observed cues, and Albert Moll in an acrimonious context (1910), refers to the "September Report" of 1904 as that "deplorable report," for he claimed to have known all along, on the basis of his own study of Hans, what was really going on. Poor Stumpf was berated for not having known, as Moll thought any worker in the area of hypnosis would, that small cues can be effective communicating agents even when unobserved by the sender or bystanders. Moll further implied

that the research program carried out by Pfungst owed much of its technique to Moll's experiments with Hans, experiments which he said had already been reported to the Psychological Society of Berlin. Ultimately Moll called Stumpf a liar (p. 458), but Stumpf survived the attack (by some 30 years, in fact) and even got in some thrusts of his own, as can be seen in some of Stumpf's supplementary sections included in this book. (When reading these sections such phrases as "obscure irresponsibles" are best read as "Albert Moll".) H. M. Johnson (1911) was right when he remarked that the name-calling aspects of the controversy did "scant credit to either of the parties" (p. 666).

Less fortunate in the matter of survival was Mr. von Osten, nor is it here quite so easy to view the controversy as somewhat quaint. Mr. von Osten, unlike scientific antagonists, had everything to lose and nothing comparable to gain by having his Hans investigated. He must indeed have been completely honest and sincere and the findings in this book grieved him deeply, nor could he accept them. Within a few months he died. (Johnson, 1911).

2. Of Pfungst's work H. M. Johnson (1911) said: "His account may be read with the interest of an exciting novel, and yet with great and lasting profit by anyone at all interested in animal behavior or in the problems of animal consciousness, whether his interest be that of the comparative psychologist, the naturalist, or the mere lover of domestic animals" (see page 633).
3. The teacher of psychology can nicely employ Pfungst's inquiry as an illustration of John Stuart Mill's major principles of scientific method as set forth in his *Logic* of 1843. Mill's

method of agreement, the observation of systematic consequences of an event, is illustrated by Pfungst's observation that Hans's cleverness followed his viewing a questioner who knew the answer to the question posed. Mill's method of difference, the more powerful principle, requires not only systematic consequences of an event but also the *absence* of these consequences in the *absence* of the event. This principle is illustrated by Pfungst's employment of blinders and of questioners who did not know the answers to their questions. Mill's method of concomitant variations, an extension of the method of difference, requires that changes in the magnitude of the consequences are systematically related to changes in the magnitude of the antecedent. This principle is illustrated by Pfungst's observation of the relationship between the questioner's distance from Hans and Hans's cleverness. (See also the discussion of Mill's methods by Boring [1954].)

4. All these clever animals were of interest to Pfungst and are cited by him in the present volume except the last, a much more recent case. Lady, the talking horse of Richmond, Virginia, was in some ways very much like Hans, only more "ambitious." She could not only count and spell but she could also locate missing objects, foretell the future, and even venture to give personal and financial advice. It was said that a friend of Lady's owner consulted Lady before making any important decisions. Like Hans, Lady attracted considerable attention, but, unlike Hans's owner, Lady's owner, a Mrs. Claudia Fonda, charged a nominal fee for her horse's services. Whatever Lady's talents were, she had a more sophisticated method of

replying to questions than Hans. Where Hans could only tap or not tap, Lady had learned to operate a special typewriter. With her nose she would flip up letters and numbers in sequence to form her brief replies. (Pictures of Lady employing her apparatus can be found in *Life* magazine of December 22, 1952, Vol. 33, No. 25.)

A thoroughly dependable informant, a research worker in another discipline (who, because of his modesty, prefers that I do not use his name), not only made me aware of the Virginia horse but observed her skill at firsthand. He and his wife owned a prize dog which had recently disappeared. The Virginia horse, reputed to be a good finder of lost objects, was consulted. My friend's wife, convinced that the dog had come to some harm, inquired of his whereabouts. The horse spelled out the word "dead." A few days later the pet turned up very much alive. My friend's interpretation, with which it is easy to agree, was that the owner of the horse sensed the questioner's conviction that the dog was dead. In some way, not at all apparent to even a keen observer, the horse's owner must have communicated to her the sequence of appropriate keyboard responses. It is one thing to find the cues that start and stop a horse's tapping. It is quite another to find the cues that lead a horse to choose one letter out of 26 and then another and another, especially when the "keys" of the typewriter are quite close together. To learn the unintentional (if it was unintentional) signaling system in this case would have provided Pfungst with another worthy challenge.

5. Pfungst's encounter with Hans was

not to be his last investigation of a clever animal. Only a few years later he was to study the conversational ability of "Don," a seven-year-old German setter. Don's "von Osten" was not an academic but a royal game warden by the name of Ebers. Don did not bark, tap, or type his answers, but rather spoke them—or so it was said. Don could "pronounce" his name, the nouns "hunger," "cake," "rest," and the verb "to have." Like Hans, Don could be asked questions in any language, though each was confined to one language in responding (taps in one case, German in the other). Pfungst showed his thoroughness once again in his investigation of Don, this time making phonographic recordings of Don's "speech." The case of Don was simpler than the case of Hans. It turned out that people simply heard him say what they expected and wanted him to say. From the phonographic recordings it became clear that if Don could speak, then he suffered from a disorder of articulation. More important, it was quickly established that Don answered all questions with answers learned in sequence. If the questions were asked in the right order, the answers, though poorly articulated, would be approximately correct. If the question order was varied, the answers were wrong. All in all, Don did not represent the same degree of challenge to Pfungst that Hans had offered or that the talking horse of Richmond might have provided if Pfungst had been there to look into it. Some additional details of the case of Don are to be found in H. M. Johnson's report (1912).

6. Let me thank Marvin Hoffman for calling this work to my attention.
7. This was a brilliant trio of investigators, all three from the University

of California at Berkeley. In accordance with the criterion of multiple mention by Boring (1948, 1950), all were major figures in the history of psychology. Compare this illustrious group to Pfungst, who, though he is mentioned in passing by Boring (1950), has been virtually unknown by American psychologists. Pfungst did indeed enjoy local fame, not only through his work with Hans but also through his writing for the popular press.

8. Let me thank E. G. Boring for calling this and other work by Johnson (and the early work of J. G. Miller) to my attention.
9. A second experiment was conducted to check on the results of this first one and to extend the generality of its findings (Rosenthal and Lawson, 1964). The learning tasks for the animals were posed within the context of Skinner boxes; half the experimenters were told that their animals were "Skinner-box dull," while the remainder were told that their animals were "Skinner-box bright." Subjects were again labeled bright and dull at random and assigned to the research groups at random. This was a more longitudinal experiment in that experimenters worked with the same animals over the course of an entire academic quarter. Seven specific experiments were performed: magazine training, operant acquisition, extinction and spontaneous recovery, secondary reinforcement, stimulus discrimination, stimulus generalization, and chaining of responses.

In this experiment, as in the earlier one, those experimenters who believed their animals were the better performers obtained significantly better performance than did the experimenters who believed their subjects

to be the poorer performers. In this study, too, experimenters who believed their subjects to be brighter rated themselves as more pleasant vis-à-vis their animals and reported that they handled them more. Although less handling of subjects occurred generally in this Skinner-box study than in the maze-learning study, there was still opportunity for handling animals in transporting them from home cage to Skinner box and then back again at the end of the day's work. At the present time, and on the basis of the two studies described, we can conclude that experimenters' expectancy may be a significant determinant of their subjects' performance when those subjects are laboratory rats. Our best guess about the "cues" involved is that they are mediated to the subject by differential handling patterns. Visual and auditory cues have not been definitively ruled out for rat subjects, but they seem less likely to play the major role they have played in the experimenter-determined performances of horses, dogs, and pigs.

10. Altogether there were 52 observers, most of whom were graduate students. For each observer a rank-order correlation was computed between the number he guessed the experimenter was after and the number that actually had been assigned that experimenter. Where by chance we would expect 26 of these 52 correlations to be positive in sign, 48 were positive, and the median correlation was $+.88$. In a replication study involving five additional experimenter-subject dyads, a smaller group of 11 observers obtained a median correlation of $+.72$ between their guesses of the experimenters' expectancy and the expectancy actually assigned.
11. The terms "kinesics" and "paralin-

guistics" refer to two of the newer areas of study within the broader field of interpersonal communication. Kinesics is the study of those body movements that have relevance for the transmission of information (for example, gestures). Paralinguistics is the study of those vocal—but not specifically linguistic—behaviors that have similar relevance (for example, tone of voice). Both areas, subsumed under the more generic area of "semiotics," have come in for recent systematic attention (Sebeok, Hayes, & Bateson, 1964).

12. The employment of expectancy control groups, including details of the analysis of the results, has been discussed elsewhere (Rosenthal, 1964a).
13. Alexander Wernicke, the physicist-philosopher from Braunschweig, thought so—especially in the case of a hypnotist's subject whose "extremely delicate sense of smell" would be mobilized to try to "read" the hypnotist's expectation. A careful experimenter, Wernicke suggested that hypnotist and subject should be separated from one another by a glass partition (Moll, 1910, p. 518).

References

- ALEXANDER, R. D., Communicative systems of animals: acoustic behavior. *Science*, 1964, 144, 713-715.
- BORING, E. G., *A History of Experimental Psychology*, 2d ed. New York: Appleton-Century-Crofts, 1950.
- BORING, E. G., The nature and history of experimental control. *Amer. J. Psychol.*, 1954, 67, 573-589.
- BORING, E. G., and BORING, MOLLIE D., Masters and pupils among the American psychologists. *Amer. J. Psychol.*, 1948, 61, 527-534. Also reprinted in E. G. Boring, *History, Psychology, and Science: Selected Papers*. New York and London: Wiley, 1963, pp. 132-139.
- BROADHURST, P. L., *The Science of Animal Behavior*. Baltimore, Md.: Penguin, 1963.
- BROWN, R., *Words and Things*. New York: Free Press of Glencoe, 1958.
- FODE, K. L., The effect of non-visual and non-verbal interaction on experimenter bias. Unpublished master's thesis, University of North Dakota, 1960.
- FOSTER, W. S., Experiments on rod-divining. *J. appl. Psychol.*, 1923, 7, 303-311.
- FRIEDMAN, PEARL, A second experiment on interviewer bias. *Sociometry*, 1942, 5, 378-379.
- GELDARD, F. A., Some neglected possibilities of communication. *Science*, 1960, 131, 1583-1588.
- GRUENBERG, B. C., *The Story of Evolution*. Princeton, N.J.: Van Nostrand, 1929.
- HANSEN, F. C. C., and LEHMANN, A., Ueber Unwillkürliches Flüstern. *Phil. Stud.*, 1895, 11, 471-530.
- JOHNSON, H. M., Review of Pfungst, O., *Clever Hans* (New York: Holt, 1911). *J. Philos., Psychol. & Scientific Methods*, 1911, 8, 663-666.
- JOHNSON, H. M., The talking dog. *Science*, 1912, 35, 749-751.
- JOHNSON, H. M., Audition and habit formation in the dog. *Behav. Monogr.*, 1913, 2, no. 3, serial no. 8.
- KELLOGG, W. N., Sonar system of the blind. *Science*, 1962, 137, 399-404.
- KENNEDY, J. L., Experiments on "unconscious whispering." *Psychol. Bull.*, 1938, 35, 526. (Abstract)
- KENNEDY, J. L., A methodological review of extra-sensory perception. *Psychol. Bull.*, 1939, 36, 59-103.
- LINDZEY, G., A note on interviewer bias. *J. appl. Psychol.*, 1951, 35, 182-184.

- MERTON, R. K., The self-fulfilling prophecy. *Antioch Rev.*, 1948, 8, 193-210.
- MILL, J. S., *A System of Logic, Ratiocinative and Inductive*. London: Longmans, 1925 (1843).
- MILLER, J. G., *Unconsciousness*. New York: Wiley, 1942.
- MOLL, A., *Hypnotism*, 4th ed. New York: Scribner's, 1898.
- MOLL, A., *Hypnotism*, 4th enlarged ed., trans. A. F. Hopkirk. New York: Scribner's, 1910; London: Walter Scott.
- ORNE, M. T., On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Amer. Psychologist*, 1962, 17, 776-783.
- POLANYI, M., *Personal Knowledge*. Chicago: University of Chicago Press, 1958.
- RAZRAN, G., Pavlov the empiricist. *Science*, 1959, 130, 916.
- ROSENTHAL, R., The effect of the experimenter on the results of psychological research. In B. Maher, ed. *Progress in Experimental Personality Research*. Vol. I. New York: Academic Press, 1964.
- ROSENTHAL, R., On the social psychology of the psychological experiment: the experimenter's hypothesis as unintended determinant of experimental results. *Amer. Scientist*, 1963, 51, 268-283.
- ROSENTHAL, R., The control of experimenter expectancy effects. Unpublished manuscript, Harvard University, 1964a.
- ROSENTHAL, R., and FODE, K. L., The effect of experimenter bias on the performance of the albino rat. *Behav. Sci.*, 1963, 8, 183-189.
- ROSENTHAL, R., and LAWSON, R., A longitudinal study of the effects of experimenter bias on the operant learning of laboratory rats. *J. Psychiat. Res.*, 1964, 2, 61-72.
- SEBEOK, T. A., HAYES, A. S., and BATESON, MARY C., eds. *Approaches to Semiotics*. The Hague: Moulton, 1964.
- STANTON, F., Further contributions at the twentieth anniversary of the psychological corporation and to honor its founder, James McKeen Cattel. *J. appl. Psychol.*, 1942a, 26, 16-17.
- STANTON, F., and BAKER, K. H., Interviewer bias and the recall of incompletely learned materials. *Sociometry*, 1942, 5, 123-134.
- STRATTON, G. M., The control of another person by obscure signs. *Psychol. Rev.*, 1921, 28, 310-314.
- WARNER, I., and RAIBLE, MILDRED, Telepathy in the psychophysical laboratory. *J. Parapsychol.*, 1937, 1, 44-51.
- WILSON, E. B., *An Introduction to Scientific Research*. New York: McGraw-Hill, 1952.
- ZIRKLE, C., Pavlov's beliefs. *Science*, 1958, 128, 1476.

INTRODUCTION

In the following article you will see the happy result of a brilliant scientist's tireless curiosity. Von Békésy treats in turn a remarkable variety of topics which have caught his interest over the years. He writes about reciprocal inhibition of opponent muscle sets and of lateral inhibition which results in Mach bands. Starting from the simple premise that all sense organs have developed out of the skin, von Békésy establishes some remarkable parallels between the delicate sensory surface on the cochlear membrane in the ear and the skin surface on the forearm! In a similar fashion von Békésy predicts that just as a human is very good at auditory "direction finding" he might be very accurate at detecting the apparent position of a stimulus presented to the two nostrils or to both sides of the tongue. Although it may take as much as a second to develop a taste sensation in response to a gustatory stimulus, von Békésy found that the surface of the tongue is sensitive to time intervals as small as 1/1000th of a second.

Not all of the arguments are very easy to follow, but the reader who works through them will gain a profound respect for this most remarkable and inquisitive Nobel laureate.

SIMILARITIES OF INHIBITION IN THE DIFFERENT SENSE ORGANS

Georg von Békésy

I would like to show in the following text the scientific framework of observation that made me interested in inhibition. Since I worked for many years in telephone engineering, the input and output system as represented in Figure 1 was, at that time, the basis of all my thinking. It is almost an electrical model of the causality principle. Unfortunately, the simple input-output model does not hold for many situations. For instance, the sound of a speaker is reflected from all the walls of a room and the listener hears two types of sounds, first the direct sounds representing the simple input-output system, and then the reflections from the walls. To localize the speaker correctly, all the reflected sounds coming from the

walls have to be suppressed. I would like to call this suppression inhibition.

Inhibition is not new. By looking at Greek statues, it is obvious that sculptors realized that it is impossible to contract all the muscles simultaneously. During movement, when one muscle is contracted its opponent is always relaxed. The relaxed muscle is the one that is inhibited by the nervous system. If we stimulate electrically both muscles simultaneously, we do not move, but have a cramp. If we look at Roman copies of Greek statues, we may feel that something is wrong and misunderstood. My impression was that the Roman copies did not understand the principle of

Reprinted with permission of the author and publisher from Georg von Békésy's "Similarities of Inhibition in the Different Sense Organs," American Psychologist, 24, 1969, pp. 707-19. Copyright © 1969 by the American Psychological Association and reproduced by permission. This was an invited lecture presented at the meeting of the American Psychological Association, September 1, 1968, in San Francisco. The research was supported in part by the American Otological Society Grant M-14, the National Science Foundation Grant GB-5768, and the National Institute of Health Grant NB-06890.

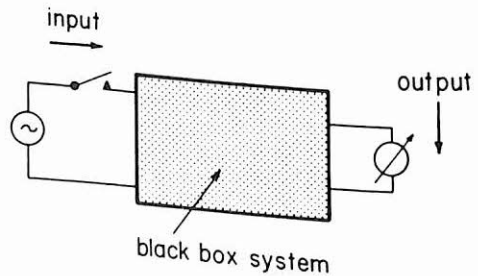


Figure 1
The input-output system. (The electrical analogy of the causality principle.)

inhibition so they represented in action large groups of muscles, contracted, which is physiologically not possible. Therefore, some muscles seem to stand out as bulges on the wrong place. It is this misunderstanding of the inhibitory process during movements that makes the Roman statues somehow unreal.

In the early Renaissance, there was no interest in the human figure. Leonardo da Vinci was the first to study anatomy, but he was so involved in anatomy that he represented most of the muscles in a completely relaxed status. There is no stimulation, and there is no inhibition whatsoever. Therefore, most of his drawings and most of his paintings look flat and really represent half-head people. The famous Mona Lisa painting shows the fascinating smile of a cadaver. This is one of the reasons why, during the Renaissance, most of the pictures showed no dynamics. In the later period, the muscles were put in action. But again, there was no discrimination between inhibited muscles and excited muscles. Many of the statues, especially the very much admired Hercules, looked, according to Cellini (the most famous goldsmith), like sacks of potatoes.

However, Eustachius, who lived from 1515 until 1574, had a new theory. His theory was that at any place that there is a joint, there has to be a muscle. I had the impression that in his later years, he found that to every joint, two muscles were attached. It was his pupil Varolius (1543–1575) who discovered in this way the stapedius muscle of the middle ear, because there was a joint between the incus and the stapes. In my opinion, the discovery of the stapedius muscle, which is the smallest muscle in the human body, is an

important discovery predicted by a theory.

I would like to call the law, that every joint is connected with a muscle, the Law of Eustachius. It is the first law in anatomy. We need many other laws of exactly that type that are able to integrate our views on anatomy and that lead to new discoveries.

The first scientist who clearly described muscle activity was Descartes (1677). As can be seen from Figure 2, he shows for the rotation of the eyeball two muscles, one of them is excited and the other one is inhibited and flat. It was evident to him that the only way to produce a rotation of the eyeball is to have this interaction between the two muscles, so that when one muscle is excited and contracted, at the same time his companion is inhibited and relaxed. This holds for

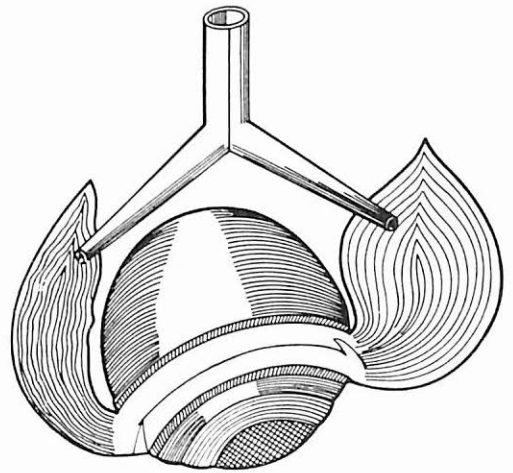


Figure 2

Eyeball with the two muscles to rotate them. The muscle on the right side is contracted, the muscle on the left side relaxed. In this situation, the eyeball is rotated counterclockwise. The upper part in the drawing represents the nerve supply. (After Descartes.)

every degree of freedom. There is a little too much seventeenth-century hydraulics in Figure 2. The muscles do not change their volume that much during contraction and the nerve tracks are not like pipe lines.

In the nineteenth and twentieth centuries, inhibition was well known. I would like to quote here Hering's (1878) theory of color vision, which is based on the assumption that there are groups of colors, which mutually inhibit each other. The work of Sherrington (1947) and Eccles (1957) is very well known, and lateral inhibition in vision was mainly described by Hartline (1949) and Ratliff (1965). Central inhibition was discovered early by the Russian scientist Sechenov (1935).

It is the discrimination between central inhibition and lateral inhibition that clarified many facts we know today. The difference can be best illustrated by some examples. It is a well-known fact that a miller hears the noise of the grinding with a very much reduced loudness. The loudness reduction is the same for all frequencies and any small event can make the loudness jump back to its original value. We think this type of inhibition is produced on the higher level of the nervous system and we would like to call it central inhibition. To say that there is a central inhibition is equivalent to the statement that the sensation magnitude is reduced. In some cases it is even difficult for the subject to say if the sensation was reduced by central inhibition or a reduction of the stimulus.

Lateral inhibition, on the other hand, is the interaction between two similar stimuli on a sense organ, particularly with a large surface area. It can produce a reduction or an increase of the local sensation magni-

tude. It produces a sharpening of the edges of a stimulus pattern and it may change the localization of a stimulus considerably.

The general concept in sensory sciences has been that inhibition is a relatively small effect that only modifies slightly the input-output system seen in Figure 1. Of course, the system had to be modified to the type shown in Figure 3 to fulfill the more com-

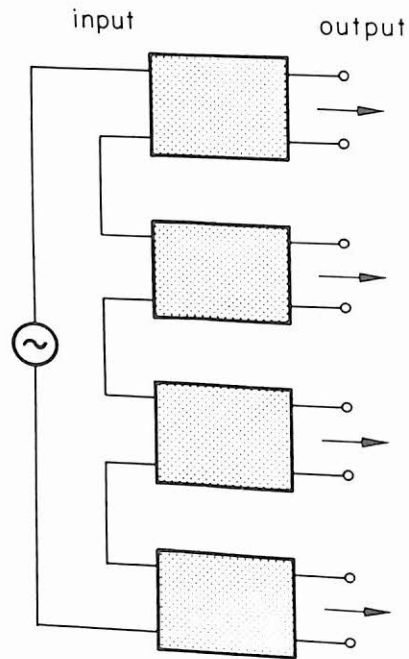


Figure 3
Improved input-output system with many outputs.

plicated requests of today where one input controls several outputs. But in general the input-output system remained and it was successful even in investigations of the brain. The reason for this is easy to understand if we look at Figure 4, a drawing made by Aerospace Medical Research Laboratory. It shows the interconnections of different parts of the nervous

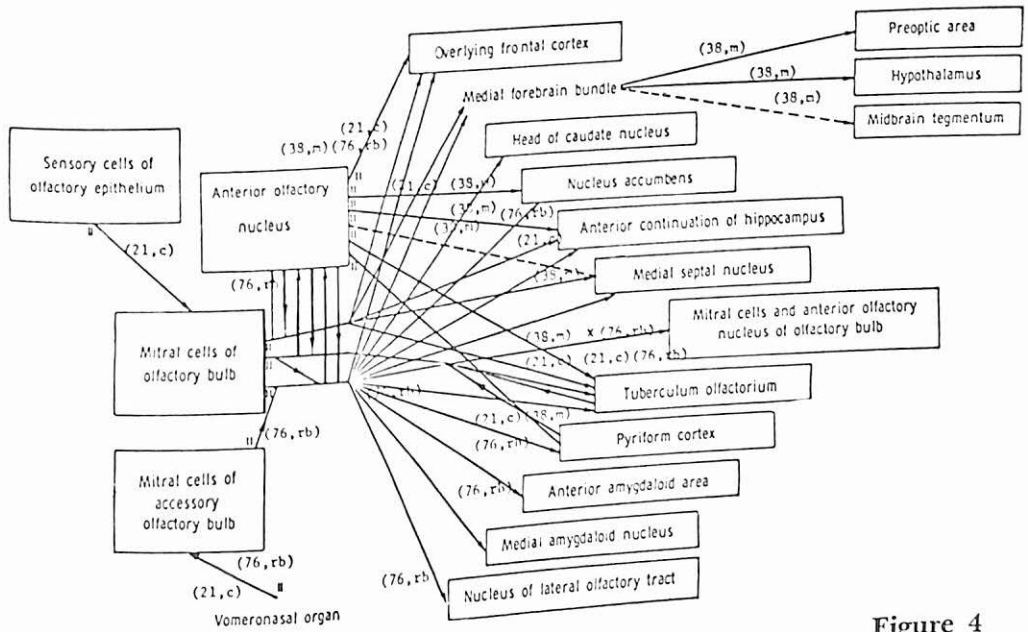


Figure 4

Interconnections of the brain that show a similarity to input-output system.
(From A. Rapoport & Co-workers, The Mammalian Central Nervous System
as a Network, 1968, mimeo.)

system in the brain. It can be seen that there are many interconnections that are very similar to the input-output system. They are connections from one central part to other parts. It is in this type of nervous connections where electrophysiology is successful, because it has all the possibilities of investigating an input and an output.

But we do not have to look around too much to be aware of the fact that this type of input-output system can be used to explain only some aspects of our sensation. Some of the problems of modern rocketry illustrate the limits of the input-output system as an explanatory model. When we aim a missile at Mars, for example, in addition to the initial aiming after it is already on its way, we constantly correct its position and direction, using

the data we get fed back as the rocket approaches the planet. The important progress of the past years in technology is that we are able to correct our aim continuously by the process of feedback. Similarly, it is quite obvious that if we want to move our finger with a high speed to a certain point, we do not use an input-output system with one single adjustment. On the contrary, we try to get a continuous feedback from our arm and from the finger, and compare the hand position with the position of the object we want to touch.

It is exactly the same procedure that is used today in modern aiming techniques. Naturally, if we have a very simple process, we would always prefer an input-output method. But, if we have to achieve something complicated, then it is obvious that the

system with the feedback is much more effective and much more precise. There was a time when in mass production all the adjusting was done once. In situations where the amount of material used was not important, this method is still used. But today, for the more complicated productions, we have to use a feedback system that will readjust the manufacturing process if anything is wrong with the end product.

In spite of the fact that in manufacturing more and more feedback systems are used, there are some very important fields where the resistance against the introduction of feedback considerations is very high. For instance, economists still insist on a five-year plan, a strict application of an input-output concept. It is a good system for the start but during later development it becomes questionable. The education program of a university with 20,000 students handled as a five-year plan without feedback can be quite a risk. Maybe the worst applications of the input-output concept are some military war projects that are maintained until there is no more input left.

If we use however a feedback system, immediately the question comes up, what should we feed back? Any electronic device will put up a great deal of information that we might not want to feed back, if it is unimportant for our purpose. Therefore, we need a selective feedback. This can be done by filter action, etc., and this filter action is the sort of inhibition that I would like to describe. We can see already from this simple description that feedback and inhibition have to be applied together.

Figure 5 shows again a drawing made by the Aerospace Medical Research Laboratory. It shows another

part of the nervous system with many feedback lines. This is the part in the brain in which corrections are produced, and I think it is the more important one. It is this type of network that is found mainly also in the microstructures of the nervous system. We immediately realize that in this type of interconnection, simple input-output measurements are difficult to make. But, if we want to go ahead with electrophysiology in the way we are doing it now, the only possibility is the use of simultaneous multi-electrode recording.

About two or three decades ago, there were very few known interconnections in the brain between the left side and the right side of the auditory tract. I almost had the feeling that some of the connections were missing and therefore I decided to wait. In the meantime, the whole picture changed completely. Today, we have so many interconnections, as you can see from Figure 5, that it is impossible to decide which one will play the role in direction hearing, to mention one example. It is almost so that every point of the brain seems to be interconnected with every other point. Unfortunately, neuroanatomy is at the moment still mainly descriptive rather than functional.

Similarities Between the Different Sense Organs

If we look at any handbook of physiology or psychology, it is surprising how neatly the different sense organs are separated into different chapters and each one is treated, in general, by a different author. Unfortunately, I find this disturbing because it creates an artificial separation barrier between different parts of physiology where we would expect

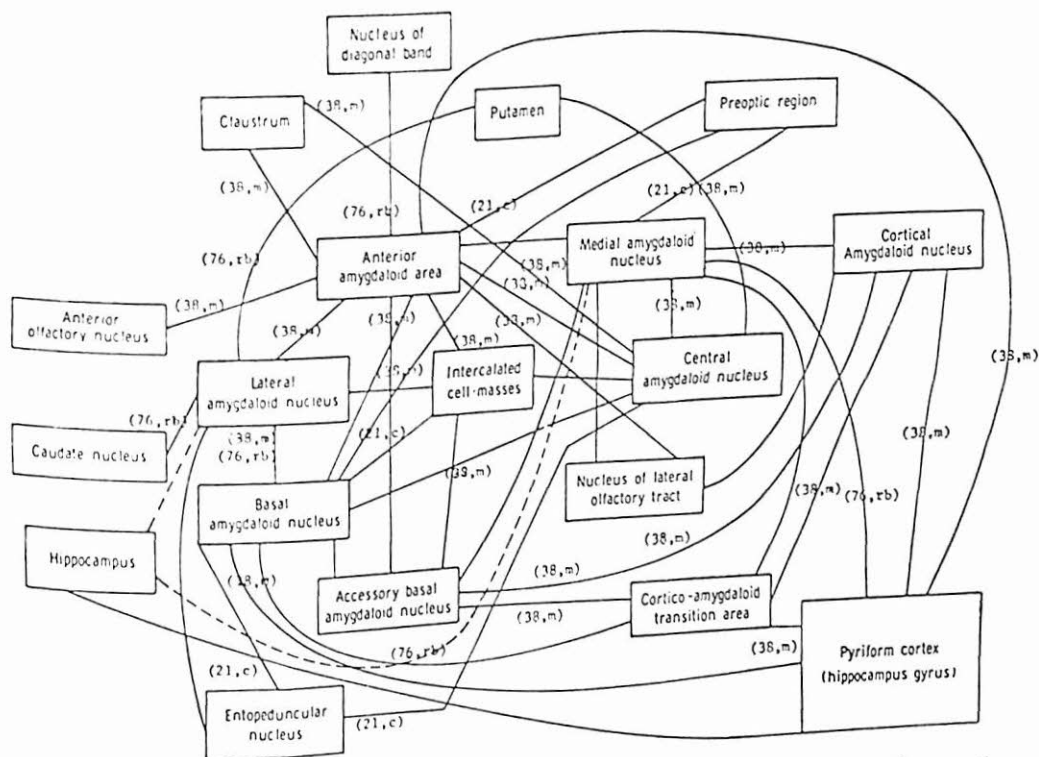


Figure 5

Interconnections of the brain that show mainly a feedback system. (From A. Rapoport & Co-workers, *The Mammalian Central Nervous System as a Network*, 1968, mimeo.)

some cross-correlations. After all, we know that all the sense organs develop from the skin. They are surely modified during evolution, but there must be some remaining similarities. From the point of view of experimental work, any similarity is of outstanding importance since there are many experiments we cannot do on one sense organ, which can be done easily on another one. For instance, it is difficult to do experiments on the basilar membrane of the cochlea, which could be done very easily on the surface of the human skin. Therefore, it is useful at least to do the preliminary experiments on the human skin, then to extrapolate what

we may expect in the cochlea. There are other experiments that are simple on the retina but are almost impossible to repeat on the basilar membrane.

If we look at Figure 5, the interconnections are so complicated that the question arises as to how can we start to do any experiments or any research in a field that is so difficult. As a beginning, we always can do one thing: we can investigate to see if this complication is about the same for all the sense organs. If the pattern is completely different for every sense organ, I definitely would give up research in sensory sciences. But, as it turns out, the patterns are similar,

and there is a certain hope that by combining the different sense organs, we might make some progress. Furthermore, I think the investigation of the similarities between the sense organs is a new avenue.

From the engineering point of view, it would be most convenient to consider the whole sensory system as simply consisting of different receptors, adapted to the different stimuli in nature. The difference between the sense organs would then be in the end organs, which would then be the transformers of the physical stimuli. And after this is done we would prefer that all subsequent neural messages and their interactions should be similar. This assumption, to put all the differences between the sense organs into the periphery and make the rest of the higher nervous level similar, actually seems to be quite fruitful. It immediately points to the fact that similarities should be expected mainly in the processing of the nervous information.

On the basis of some preliminary experiments, this conclusion seems to be correct. The so-called Mach band may be used as an illustration. As shown in the top drawing of Figure 6, we can produce on the retina a light distribution that increases from zero up to a certain constant value, then drops continuously down to zero. This luminance (physical) distribution represents in this case a trapezoid. We will expect that the brightness (psychological) distribution would be very similar to a trapezoid also, which should show only one maximum with a flat top. But, as illustrated in the drawing below, this does not occur. We have two maxima. These two maxima are called white Mach bands. They were discovered 100 years ago by Mach (1865-1868) and they indi-

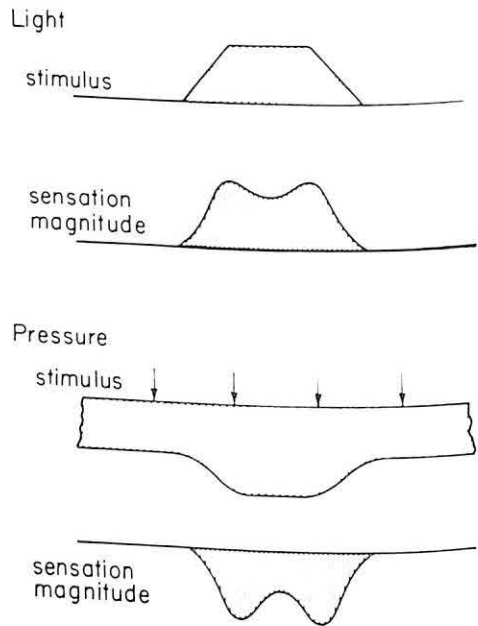


Figure 6

Luminance distribution along the surface of the retina that has a trapezoidal form, as shown in the top drawing, produces a brightness distribution that shows two maxima. The two lower drawings indicate that a cardboard with a trapezoidal cut pressed against the surface of the skin will produce a pressure sensation similar to that of the brightness distribution with two maxima in spite of the fact that the original pressure stimulus has only one maximum.

cate that the distribution of the sensation magnitude does not completely correspond to the stimulus distribution, rather the sensation emphasizes discontinuities in the stimulus distribution. As we can see from the second drawing, the corners particularly are small and it is very easy to demonstrate with simple equipment.

It is not difficult to conduct the analog experiment for the pressure sensation on the skin of the arm. To produce a certain pressure distribu-

tion we can cut out from a cardboard the shape seen in the third drawing. Again, the pressure distribution will be in the form of a trapezoid. We can test the magnitude of the pressure distribution by measuring under a microscope the indentations of the skin, and it turns out that it follows exactly the pattern of the trapezoid. Any observer will immediately say he feels practically only the corners as it is shown in the lowest drawing of Figure 6. We have the Mach bands not only in the retina, therefore, but we have them also on the surface of the skin, a similarity already pointed out in 1928. Since that time, I have been interested in discovering more and more similarities between the different sensory perceptions.

It is well known that the direction of an incoming sound can be well determined by the use of both ears. The direction localization is especially very sharp for clicks. If the sound is in front of us, the clicks reach both ears at the same time. But, if the sound is a little on the side, then one ear will receive the sound earlier than the other. In this case we localize the sound source into the ear that receives the sound first. The localization phenomena are so strong for a certain time difference that we have the impression that no sound reaches the ear with the delayed click at all. It is an almost complete inhibition. The time difference, between clicks to either ear which is necessary to localize the sound source completely on one side, is very small. It is of the order of one millisecond.

Since this is an inhibitory phenomenon and the time difference is so short, it is interesting to find out if in other sense organs a similar short time delay between two points of stimulation like the left and the right

ear can produce a similar phenomenon. If we then explore the taste sensation, we find the same time difference of one millisecond produces a taste sensation on the side of the earlier stimulus. We know from previous experiments that for the ear we need about 0.2 seconds for the full development of the loudness sensation. But the development of the taste sensation takes more than a second. Therefore, it was surprising to find that even for a sense organ with such a slow-acting receptor, the time difference of one millisecond between a two-point stimulation could be recognized.

The measurements for the taste experiments were extremely difficult and required a complicated equipment because we had to exchange the fluid on the surface of the tongue in a very short time interval. We used a plastic plate as is shown in Figure 7. In this plastic plate, two openings were drilled that were slit open on the surface of the plate. The plate placed in the middle of the tongue permitted a contact between the fluid and the surface of the tongue. There were two slits, one on the left and the other on the right side of the tongue. All connections to these slits have to be made extremely short and a special equipment was made to exchange the constantly streaming tap water or the physiological solution with the taste solution we wanted to investigate. The experiments indicated that the better the equipment, the shorter was the time delay necessary to change the localization of the taste sensation from the middle of the tongue to the left or to the right side. After the exchange of the water with the test solution was made as soon as possible and checked with special recording devices placed on the surface of the

and there is a certain hope that by combining the different sense organs, we might make some progress. Furthermore, I think the investigation of the similarities between the sense organs is a new avenue.

From the engineering point of view, it would be most convenient to consider the whole sensory system as simply consisting of different receptors, adapted to the different stimuli in nature. The difference between the sense organs would then be in the end organs, which would then be the transformers of the physical stimuli. And after this is done we would prefer that all subsequent neural messages and their interactions should be similar. This assumption, to put all the differences between the sense organs into the periphery and make the rest of the higher nervous level similar, actually seems to be quite fruitful. It immediately points to the fact that similarities should be expected mainly in the processing of the nervous information.

On the basis of some preliminary experiments, this conclusion seems to be correct. The so-called Mach band may be used as an illustration. As shown in the top drawing of Figure 6, we can produce on the retina a light distribution that increases from zero up to a certain constant value, then drops continuously down to zero. This luminance (physical) distribution represents in this case a trapezoid. We will expect that the brightness (psychological) distribution would be very similar to a trapezoid also, which should show only one maximum with a flat top. But, as illustrated in the drawing below, this does not occur. We have two maxima. These two maxima are called white Mach bands. They were discovered 100 years ago by Mach (1865–1868) and they indi-

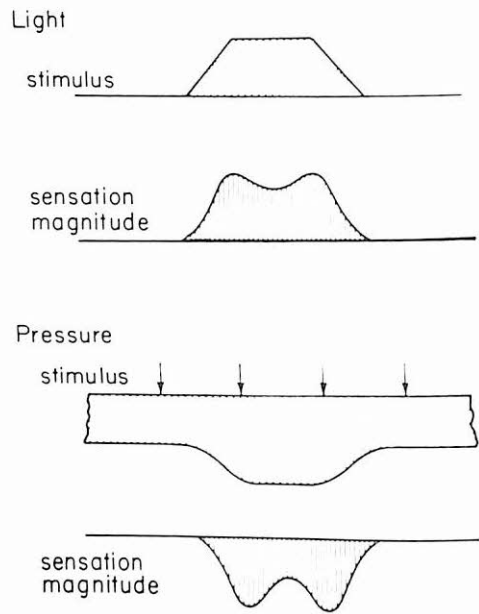


Figure 6

Luminance distribution along the surface of the retina that has a trapezoidal form, as shown in the top drawing, produces a brightness distribution that shows two maxima. The two lower drawings indicate that a cardboard with a trapezoidal cut pressed against the surface of the skin will produce a pressure sensation similar to that of the brightness distribution with two maxima in spite of the fact that the original pressure stimulus has only one maximum.

cate that the distribution of the sensation magnitude does not completely correspond to the stimulus distribution, rather the sensation emphasizes discontinuities in the stimulus distribution. As we can see from the second drawing, the corners particularly are emphasized. This phenomenon is not small and it is very easy to demonstrate with simple equipment.

It is not difficult to conduct the analog experiment for the pressure sensation on the skin of the arm. To produce a certain pressure distribu-

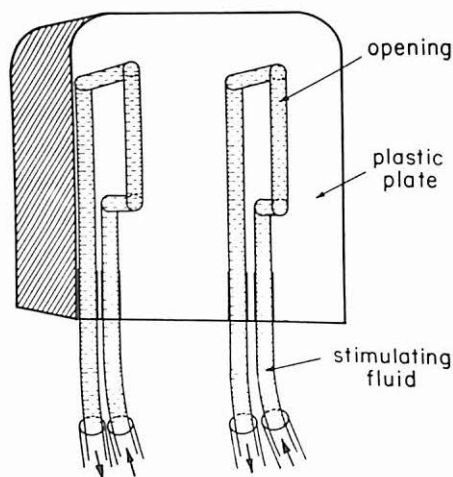
tion we can cut out from a cardboard the shape seen in the third drawing. Again, the pressure distribution will be in the form of a trapezoid. We can test the magnitude of the pressure distribution by measuring under a microscope the indentations of the skin, and it turns out that it follows exactly the pattern of the trapezoid. Any observer will immediately say he feels practically only the corners as it is shown in the lowest drawing of Figure 6. We have the Mach bands not only in the retina, therefore, but we have them also on the surface of the skin, a similarity already pointed out in 1928. Since that time, I have been interested in discovering more and more similarities between the different sensory perceptions.

It is well known that the direction of an incoming sound can be well determined by the use of both ears. The direction localization is especially very sharp for clicks. If the sound is in front of us, the clicks reach both ears at the same time. But, if the sound is a little on the side, then one ear will receive the sound earlier than the other. In this case we localize the sound source into the ear that receives the sound first. The localization phenomena are so strong for a certain time difference that we have the impression that no sound reaches the ear with the delayed click at all. It is an almost complete inhibition. The time difference, between clicks to either ear which is necessary to localize the sound source completely on one side, is very small. It is of the order of one millisecond.

Since this is an inhibitory phenomenon and the time difference is so short, it is interesting to find out if in other sense organs a similar short time delay between two points of stimulation like the left and the right

ear can produce a similar phenomenon. If we then explore the taste sensation, we find the same time difference of one millisecond produces a taste sensation on the side of the earlier stimulus. We know from previous experiments that for the ear we need about 0.2 seconds for the full development of the loudness sensation. But the development of the taste sensation takes more than a second. Therefore, it was surprising to find that even for a sense organ with such a slow-acting receptor, the time difference of one millisecond between a two-point stimulation could be recognized.

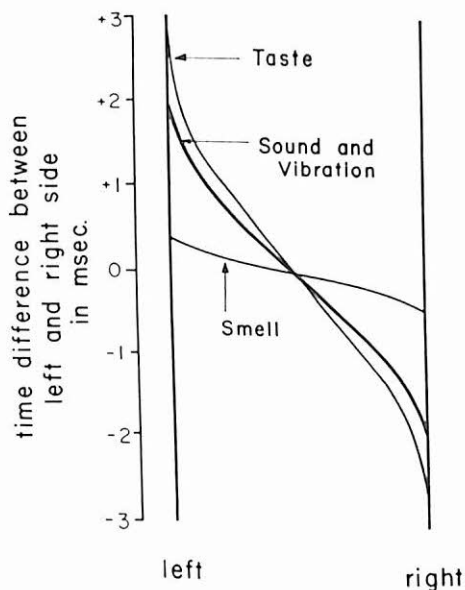
The measurements for the taste experiments were extremely difficult and required a complicated equipment because we had to exchange the fluid on the surface of the tongue in a very short time interval. We used a plastic plate as is shown in Figure 7. In this plastic plate, two openings were drilled that were slit open on the surface of the plate. The plate placed in the middle of the tongue permitted a contact between the fluid and the surface of the tongue. There were two slits, one on the left and the other on the right side of the tongue. All connections to these slits have to be made extremely short and a special equipment was made to exchange the constantly streaming tap water or the physiological solution with the taste solution we wanted to investigate. The experiments indicated that the better the equipment, the shorter was the time delay necessary to change the localization of the taste sensation from the middle of the tongue to the left or to the right side. After the exchange of the water with the test solution was made as soon as possible and checked with special recording devices placed on the surface of the

**Figure 7**

A plastic block showing the tubes through which, for taste stimuli, a fluid can be conducted. (The surface of the plastic tube is open so that the plastic tube placed on the surface of the tongue will bring it in contact with the fluid stimuli, one on the left side and the other on the right side.)

tongue, the necessary time difference to shift the taste sensation from the middle to the side was about the same as it was for sound. (See Figure 8.)

Next we explored the sensation of smell. For this experiment we introduced a tube in one nostril and a second tube in the nostril of the other side. If in the constant airflow of the two tubes an odorous substance was injected exactly at the same time, the observer located the smell in the middle plane of his nose, but if it had a time delay on one side, then he located the smell either into the left or the right nostril. This phenomenon is very pronounced and it is very similar to the general features of direction hearing. As is shown in Figure 8, the time delay necessary to displace the locus of an odorous substance from

**Figure 3**

If two stimuli are presented, and if there is no time delay, the sensation is localized into the middle between the two stimuli. (If the time delay is introduced, the sensation moves laterally. For vibrations, taste, and smell, the time delay necessary is about of the same order of magnitude.)

the middle to one side is about of the order of one-half millisecond.

There are even more complicated and interesting analogies between hearing and smelling if we watch a dog following a trail. His head movements during the maneuver are exactly the same as a dog will do if he traces a sound source. It is a continuous movement of the head to the left and to the right after which he tries to center into the correct direction.

From Figure 8 we can see that if the sensation magnitudes on both sides of the head are the same and if they stimulate the sense organs exactly at the same time, we will localize the sensation in the median plane. But,

if the time delay of about one milli-second is introduced on one side, the sensation will shift to the side that received the stimulus first. For all different sensations like sound, taste, smell, and vibration, we found about the same time delay. This seems to indicate that the localization phenomenon involves a nervous process of inhibition that is very similar for all these different sensations.

This analogy holds even though different sense organs are innervated by different numbers of nerve fibers. For vibratory sensations, we may use about 1,000 nerve fibers, but for hearing 30,000. Olfaction has about one million nerve fibers.

In localizing a sensation, the apparent spatial size of the sensation plays an important role. It is well known in direction hearing that the apparent size of a click is small. Furthermore, the sharper the click is, the smaller the apparent size. On the other hand, a soft low-frequency tone appears much larger. We would expect then, on the basis of this difference in apparent size, that a low-frequency tone would be less precisely localized than a click. This is in fact what our experiments showed.

It is easy to observe further, that in audition, there is very little change of the apparent size of a sound image during its movement from the midline to the side. This is illustrated in Figure 9. However, it is not so for vibration sensations produced on the skin. If the two vibrators that are interacting are far away on the arm, the vibratory sensation in the middle between them is diffused. If the vibratory image is on one side, then we have a very well-defined image of the vibratory sensation that is not much larger than the vibrating surface that touches the skin. It is interesting to

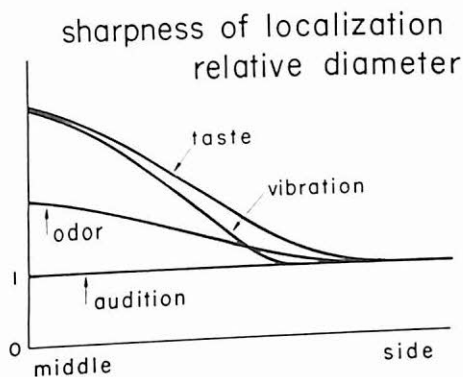


Figure 9

For audition, the apparent magnitude of the sound image does not change if the localized sound is moved from the middle to the side by introducing a time delay. (Other sense organs do not show this constancy.)

notice that if the two vibrators are placed close to each other, then the uncertainty becomes less pronounced. If we bring them very close together, we have practically no change in the extension of the vibratory sensation during its movement from the middle to the side. The situation for the taste sensation and olfaction is also shown in Figure 9. Here again, the closer the two stimuli are, the less is the change in the extension of the sensation area during movement. We may assume that if the two stimuli are closer together, then there is more neural interaction between the two points. We can conclude from the experiments that the more interaction is present, the less is the change of the sensation area during the movement from the middle to the side.

Summing up, it seems that the localization phenomena produced by time differences in the different senses are surprisingly similar. Localization in all the senses is produced by short

time delays below one millisecond, despite the well-known fact that the time necessary to fully develop a sensation magnitude always takes more than 200 milliseconds.

Oscillation of the Sensation Magnitude

An effective way to show that sensations are not produced by a simple input-output system is that in many sensations, oscillations can be observed. If we switch on a continuous stimulus, this will start a set of oscillations that decay under normal conditions of the nervous system. This phenomenon is known for vision. Especially, it can happen under certain conditions on the periphery of the visual field, that we see a continuous fluctuation of brightness. It is a yellow flicker around the whole periphery. The frequency of the flicker is constant and it is not easy to change anything about it. The flicker is an internal process that is not modified by external stimuli, for instance, by looking at a bright light or going into a dark room.

It is obvious that in any system that has a feedback, we expect oscillation. Most people of the older generation remember that the biggest problem of amplifiers was to make them free of oscillation. Fortunately, there were some rules that permitted the damping of oscillations in a simple system like an amplifier. But in a complex system as seen in Figure 5, to my knowledge there are no rules available. I was always interested in the question as to whether it is possible to have in such a complex system two different oscillations simultaneously. I had once several hundreds of telephone relays switched in a random way together. They all had their

energy supply and in some cases a time delay produced by a condenser parallel to the relay coil. If one of the relays was switched on, then in general a large part of the relays also switched to a different position. But in most cases no oscillations were observed. Only by introducing long feedback lines was it possible to produce oscillation. These oscillations all had the same frequency. There was one type of oscillation that seemed interesting since it represented a sort of sweeping of the relay positions from one side of the whole set of relays to the other side in a periodic way. To produce this type of behavior, it is necessary to have very large numbers of relays with definite phase difference between one side of the set and the other one. If the traveling time is too short, no such sweeping will occur. I was never able to find the conditions under which such a sweeping oscillation was produced.

The last time I had the opportunity to see such a sweeping oscillation was at Cambridge in William James Hall. William James Hall consists of 15 floors and had three, more or less, automatic elevators. I was always surprised how long I had to wait for an elevator. For weekdays the explanation was that the elevators were busy. But, when I went to William James Hall on a Sunday, 7:30 A.M., it was obvious that nobody was in the building and still it took seven and more minutes until one elevator arrived. By listening to the doors, it was possible to hear the elevators running up and down, all three of them, and then suddenly like magic, all three doors opened simultaneously. After having waited so long, I was always afraid to step in any of these elevators. This type of elevator behavior is

called, in elevator language, "hunting." It is a very similar type of sweeping movement like we found in many systems. Unfortunately, again it was not possible to find out what the real conditions for this situation were. Naturally, by changing the resting location of the elevator, etc., the phenomenon could be eliminated.

To test experimentally if the switching on of a constant stimulus can produce oscillations in the sensation magnitude, the method illustrated in Figure 10 of two separate stimuli of

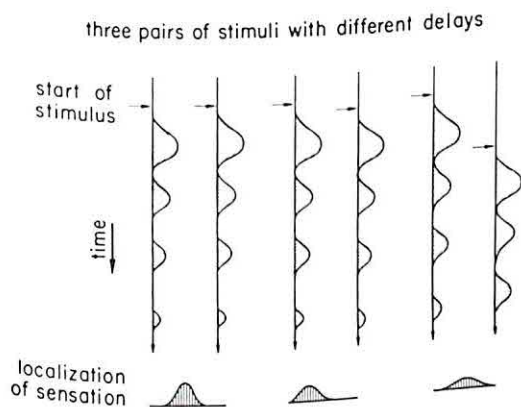


Figure 10

Sensation localization can be used to determine if a constant stimulus may produce fluctuations in the sensation magnitude. (From G. v. Békésy, 1967.)

equal sensation magnitude can be used. If both stimuli start exactly at the same moment, we may get sensation magnitude changes with time as illustrated in the drawings. Since both sensations are exactly the same, we will locate the sensation in the middle between the two stimulators, as indicated for the stimulus pair on the left side. In this case again, the sensation will be localized in the middle position. Naturally, the sensation magnitude will not be exactly the same for the second and third maximum.

Therefore, the position in the middle is not as precise as it was when both stimuli were presented without any time delay.

We did this experiment on the surface of the tongue by using two concentric electrodes on both sides of the tongue. The current through these electrodes was adjusted so that a taste sensation was produced in the middle of the tongue if no time difference was present. If a continuously increasing time delay is introduced, the localization of the taste moves repeatedly from one side to the other. This indicates that for the taste sensation, stimulated electrically, there is an oscillation of sensation magnitude present. The frequency of the oscillation can be determined from the time delay necessary to produce a repetition. We could repeat the same observations for stimulations with a salt solution, and find about the same period as found for the electrical stimulation.

Inhibition Is Not Small

In general, it has been assumed that inhibition plays a very unimportant role in daily life and in experiments with sense organs. Some forty years ago, I started to point out that between a stimulus distribution on a sense organ with a large surface and the observed sensation distribution, there is a very large difference under some conditions.

In Figure 11, you can see a vibrator vibrating with a frequency of 50 cycles per second. It is touching the surface of the skin on the lower arm. With stroboscopic illumination, we can see a traveling wave along the arm, and we can measure its wavelength and decay in amplitude and, therefore, we would expect to feel vibrations along

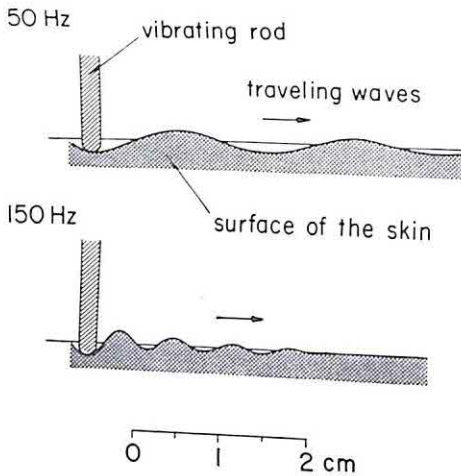


Figure 11

A vibrating rod produces traveling waves along the surface of the skin.

the whole arm. But if we ask an observer, he reports that the vibratory sensation extends about one centimeter around the vibrator. The more remote sections of the traveling wave are completely inhibited. The inhibition is produced partially by the time delays of the successive sections of the traveling wave. This situation does not change even when we show to the observer his whole arm vibrating. The phenomenon is even more dramatic if the observer stands on a vibrating table. His whole body is seen vibrating and he still insists that he feels only vibration on his feet.

In the field of hearing, it is especially easy to devise experiments that show a large amount of inhibition. When we listen to a speaker in a room, we localize all the sounds reflected from the walls as coming from the place of the speaker. This process of inhibition is a very important phenomenon. If we place one microphone about one foot away from the speaker, and a second microphone

about five feet away, the sound pattern of these two microphones is completely different, as can be seen in Figure 12. In the lower recording the reflections from the wall distort the original sound pressure pattern close to the speaker. In spite of the large difference between the sound pressures, a listener does not hear a difference when he moves his head from one microphone to the other, since in binaural hearing the reflections from the walls are inhibited.

The knowledge of these wall reflections and the inhibitory processes concerning localization is of utmost importance for the acoustics in concert halls and theaters. It can make the difference between a good or a bad concert hall.

The Neural Unit

There are many stimulus patterns by which a sense organ with a large surface area can be excited. It is of interest to find a general approach to such stimulation.

Mach, almost 100 years ago, developed a formula for transforming a

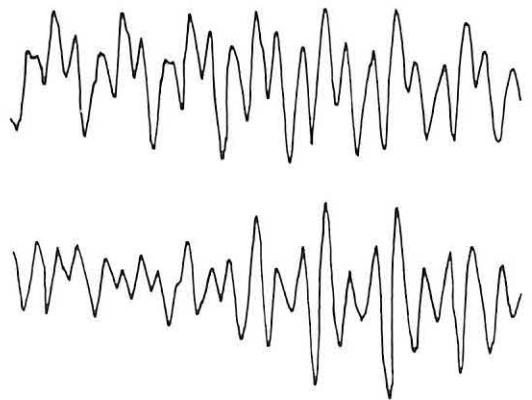


Figure 12

The sound pressure can be very different on two places and still the ear will not recognize any difference.

stimulus distribution along the retina into the sensation magnitude distribution, that is, for transforming the luminance pattern on the retina into the brightness pattern in sensation. His formula contains, as a main distorting factor, the second derivative of the luminance function. Mach's formula describes very well the appearance of the white and the dark bands. It is in agreement with the observations that the smaller the curvature of the luminance distribution, the larger the brightness of the dark or white band. If the curvature is positive, we have a white Mach band, and if the curvature is negative, we have a black band. I became interested in finding some physiological background for this phenomena, in addition to the purely mathematical description. To test this question, the simplest approach seemed the investigation of how two stimuli interact with each other if presented to a sense organ with a large surface. The sensation produced by one single point stimulus has a certain spread, but it is not large. If we have two stimuli that are close together, the magnitude of the sensation increases, and the width of the spread also. We can go farther and farther apart with the two stimuli, and there is a certain distance when the sensation becomes as wide as the distance of the two stimuli. We have only then to increase the distance just a little more and, at that moment, we have the peculiar result that the magnitude of the combined sensation suddenly drops (Békésy, 1960). The more precise the two sensation magnitudes are adjusted to be equal, the larger is the drop. If we increase the distance, the two sensations will fall apart. The simple fact that two stimuli can give a smaller sensation magnitude than one means

that there is a lateral inhibition. This phenomenon was tested for pressure on the skin and for taste.

It is possible to describe these findings with the help of the drawings shown in Figure 13. These drawings

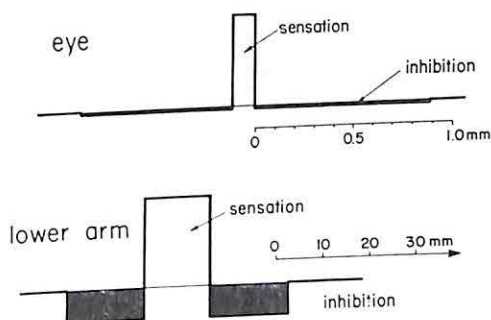


Figure 13

The area of sensation and inhibition for vision and the skin of the lower arm.

indicate that any stimulus of a small area will produce an area of excitation of a certain width, and simultaneously a surrounding area of inhibition where the sensation magnitude produced by a second stimulus is decreased. The combination of both areas (excitation and inhibition) may be called an inhibitory unit for one stimulus. As a first approach, the unit is made out of rectangles, and the dimensions of the width of the sensation area and the width of the inhibited area for vision and skin sensations are shown.

In Figure 14, we have a stimulus distribution that has on the left side a constant stimulus and then going to the right side increases linearly to a certain value that remains constant again. In order to analyze this distribution in terms of the above inhibitory unit, we use the following procedure. The stimulus distribution is cut in small vertical segments of the

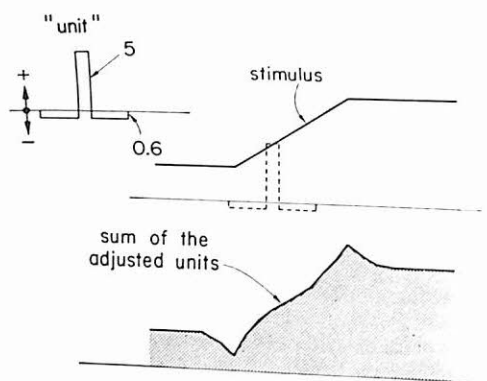


Figure 14
Method to determine the difference between the stimulus distribution and the sensation distribution. (From G. v. Békésy, 1960.)

width of the excitation. The stimulus magnitude is multiplied by the excitation magnitude of the unit and they are plotted all along the stimulus distribution curve. The same thing is done for the inhibitory areas, and after the two curves are drawn, we subtract from the excitation magnitude distribution on the inhibitory magnitude distribution. As can be seen in the lower drawing of Figure 14, the resulting sensation magnitude curve shows a well-developed overshoot and undershoot. The width of this overshoot and undershoot is given by the width of the inhibitory area of the unit. It was found that for many stimulus distributions, this type of geometrical procedure is successful. The geometrical addition is different than the pure mathematical formula of Mach. For instance, for a very sharp corner, the curvature becomes infinite and therefore the mathematical overshoot and undershoot should be extremely high also. In the geometrical analysis, the overshoot has always a certain width but not an infinite height.

Displacement of the Sensation Maximum

Lateral inhibition can produce a difference between stimulus distribution and sensation distribution that is quite surprising. We cut out from a black sheet an opening as shown on the left side of Figure 15. This open-

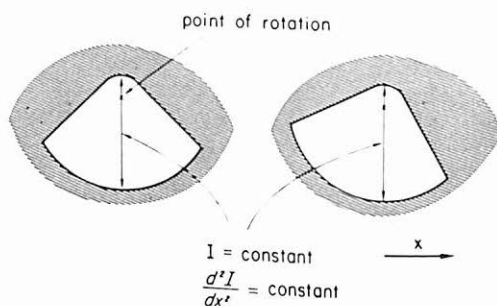


Figure 15

Masks to produce a luminance distribution along the surface of the retina for which in a certain region the luminance and its second derivative are constant.

ing has at the top, and on the lower edge, a constant curvature. It can be rotated around the center of the curvatures (marked by a cross). A rotation of the opening around its center will not change the stimulus magnitude or its second derivative.

We can plot the height of the opening in a curve as it is shown in the upper drawing of Figure 16 (Békésy, 1968). In the represented situation, the opening is symmetrical to the vertical line. From the stimulus distribution, we can calculate the sensation magnitude distribution by using the inhibitory unit. In this symmetrical situation, the stimulus has a maximum that falls together with the maximum of the sensation. We cannot see any reason why this should change if we rotate the opening in Figure 15 around the center point,

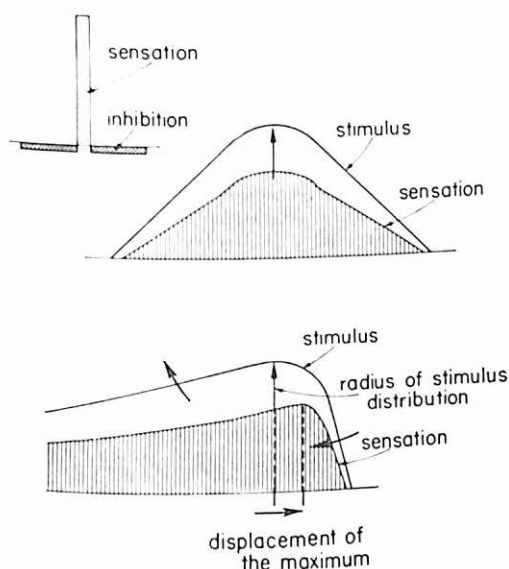


Figure 16

A stimulus distribution that keeps its maximum and second derivative constant can still produce a displacement of the sensation maximum if the slopes on the side change. (From G. v. Békésy, 1968.)

because neither the stimulus magnitude nor the second derivative near the point of the maximum will change. But using again our geometrical method, we find that for the sensation distribution, the maximum is quite away from the maximum of the stimulus distribution, as shown by the shaded area in the lower drawing of Figure 16.

The question is, do we find the similar type of the shift of the sensation magnitude relative to the stimulus maximum in actual observations? To test this in vision we use an equipment that transforms the light of the openings in Figure 15 linearly into a luminance distribution (Békésy, 1968). Experiments showed that for the symmetrical position of the opening the stimulus maximum coincided

with the brightness maximum. However, if we rotated the opening, the brightness maximum shifts to the side where the stimulus falls off. During rotation, we have a feeling that the maximum of the sensation moves into the area where the stimulus is reduced, a paradox phenomena but in agreement with Figure 16.

We repeated this experiment for pressure and heat sensations of the skin with the same results. I was always interested in the pain sensation produced by certain inflammatory diseases. Comparing the findings of their autopsies with the descriptions patients gave during their admittance to the hospital, I found in several cases, that the pain localization was not localized at the middle of the inflammatory process, but near the edge of the inflammation.

After the war, it was quite common that the roots of the teeth on the lower jaw were inflamed. In general, the middle part of the teeth had the maximum inflammation and it dropped on different sides with a different slope. This could be demonstrated easily by X-ray pictures. But the patients, in general, did not localize the pain at the maximum of the inflammatory process, determined by the X-ray pictures, but more toward the edge, when the distribution of the inflammation did not drop equally on both sides. In this case, there was a danger to have the wrong tooth pulled.

References

- BÉKÉSY, G. v., Neural inhibitory units of the eye and skin: Quantitative description of contrast phenomena. *Journal of the Optical Society of America*, 1960, 50, 1060-1070.

- BÉKÉSY, G. v., *Sensory inhibition*. Princeton: Princeton University Press, 1967.
- BÉKÉSY, G. v., Location of maxima and minima in sensation patterns influenced by lateral inhibition. *Journal of Applied Physiology*, 1968, 25, 200-206.
- DESCARTES, R., *L'Homme*. (2nd ed.) Paris: Charles Angot, 1677. Esp. pp. 15-18.
- ECCLES, J. C., *The physiology of nerve cells*. Baltimore: Johns Hopkins Press, 1957.
- HARTLINE, H. K., Inhibition of activity of visual receptors by illuminating nearby retinal areas in the Limulus eye. *Federal Proceedings*, 1949, 8, 69.
- HARTLINE, H. K., RATLIFF, F., & MILLER, W. H., Neural interaction in the eye and the integration of receptor activity. *Annals of the New York Academy of Sciences*, 1958, 74, 210-222.
- HARTLINE, H. K., RATLIFF, F., & MILLER, W. H., Inhibitory interaction in the retina and its significance in vision. In E. Florey (Ed.), *Nervous inhibition*. London: Pergamon, 1961.
- HARTLINE, H. K., WAGNER, H. G., & RATLIFF, F., Inhibition in the eye of limulus. *Journal of General Physiology*, 1956, 39, 651-673.
- HERING, E., *Zur Lehre vom Lichtsinne* (Vienna: Gerolds Sohn, 1878). (Trans. by Leo M. Hurvich & Dorothea Jameson.) Cambridge: Harvard University Press, 1964.
- MACH, E., Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut. I. *Akademie der Wissenschaften in Wien, Sitzungsberichte, math.-nat. Cl.*, 1865, 52, Abt. 2, 303-322. Über die physiologische Wirkung räumlich vertheilter Lichtreize. *Ibid.*, 1866, 54, Abt. 2, 393-408, and 1868, 57, Abt. 2, 11-19.
- RATLIFF, F., *Mach bands: Quantitative studies on neural networks in the retina*. San Francisco: Holden-Day, 1965.
- SECHENOV, I., *Selected works*. Moscow-Leningrad: State Publishing House for Biological and Medical Literature, 1935.
- SHERRINGTON, C., *The integrative action of the nervous system*. New Haven: Yale University Press, 1947.

INTRODUCTION

Many psychologists interested in establishing a *science* of human behavior have felt bound to restrict their attention to easily measurable overt behavior which was available to any outside observer. Although there has always been great interest in what goes on inside the head, the study of "mental events" has been left mostly to the poets and the philosophers because the scientists have considered them totally inaccessible.

In the following article Sternberg describes a beautiful series of experiments in which he has found some elegant simplicity in the way a man searches his short-term memory. Suppose that a psychologist is going to show you a series of single digits and that your task is as follows: If the digit is a member of the following set (1, 3, 5, 6, 8) you should make a response with your right hand, but when the digit is not in that set then you should make a response with the left hand. Now suppose that the digit 7 is presented as a test digit. Obviously you would respond almost immediately with your left hand since 7 is not a member of the above set. But do you suppose it would be possible to find out exactly what is going on in the very brief interval between the presentation of the digit 7 and your response with your left hand? Sternberg has demonstrated that it is indeed possible and that one can specify in most amazing detail the kinds of mental processes which go on and the length of time required for many of them. Before reading this paper you might try to puzzle for a while about how you would go about extracting such apparently inaccessible information.

MEMORY-SCANNING: MENTAL PROCESSES REVEALED BY REACTION-TIME EXPERIMENTS¹

Saul Sternberg

Abstract. *Eight experiments on the retrieval of information from human memory are informally reviewed in this paper. In the interpretation of these experiments inferences are made from the structure of reaction-time (RT) data to the organization of mental processes. These inferences are based on a proposal first made by Donders (1868) that the time between stimulus and response be regarded as the sum of the durations of a series of processing stages. Donders' subtraction method depends on this stage theory, together with an assumption of pure insertion which states that a change in the subject's task can cause the insertion of an additional processing stage without altering the other stages. It was the questioning of this assumption, and the absence of any objective tests of its validity, that led to the decline of the subtraction method in the late nineteenth century. In the present paper it is shown how it can sometimes be replaced by a weaker assumption of selective influence, and how the validity of either assumption for a given experiment can be tested by determining whether the effects of experimental factors on RT are additive.*

The resulting modern version of Donders' approach has led to the discovery of two kinds of memory search that people use in the retrieval of information from short memorized lists. One is a high-speed exhaustive scanning process, used to determine the presence of an item in the list; the other is a slow self-terminating scanning process used to determine the location of an item in the list. Among the questions considered in relation to these processes are the following: To what extent is a test item preprocessed before being compared to the memorized list? Does retrieval of nonsymbolic items involve the same processes as retrieval from lists of digits? How does retrieval from "active memory" compare to retrieval from information held only in the "inactive memory"? Is retrieval from active memory influenced by how well the list has been learned? Why should a search process be exhaustive, instead of terminating as soon as the information has been found? Do recognition and recall depend on the same search process?

*We are grateful to Bell Telephone Laboratories for the figures and to Saul Sternberg, Little, Brown & Co., and Bell Telephone Laboratories for their kind permission to include this selection. Reprinted from "Memory-Scanning: Mental Processes Revealed by Reaction-Time Experiments" by Saul Sternberg in *Cognition and Affect*, John Antrobus (Ed.). Copyright © 1970 by Little, Brown & Co.*

¹Most of the research reported in this paper was supported by Bell Telephone Laboratories and conducted in its Behavioral and Statistical Research Center at Murray Hill, N.J. The work reported as Exp. 4 was done in collaboration with A. M. Treisman of the University of Oxford. I am grateful to C. S. Harris, T. K. Landauer, H. Rouanet, and R. Teghtsoonian for helpful criticisms of the manuscript, and to L. D. Harmon for discussion leading to Exp. 5. R. E. Main assisted with Exps. 4 and 5, B. Barkow with Exps. 7 and 8, and B. A. Nasto with Exps. 4, 5, 7, and 8.

One of the oldest ideas in experimental psychology is that the time between stimulus and response is occupied by a train of processes or *stages*, some being mental operations, which are so arranged that one process does not begin until the preceding one has ended. This *stage theory* implies that the reaction-time (RT) is a *sum*, composed of the durations of the stages in the series, and suggests that if one could determine the component times that add together to make up the RT, one might then be able to answer interesting questions about mental operations to which they correspond. The study of RT should therefore prove helpful to an understanding of the structure of mental activity.

The use of results from RT experiments to study stages of information processing began about a century ago with a paper, "On the Speed of Mental Processes," by F. C. Donders (1868). It was in this paper that Donders introduced the *subtraction method*—a method for analyzing the RT into its components and thereby studying the corresponding stages of processing.

1. Decomposing RT by the Subtraction Method

To use the subtraction method one constructs two different tasks in which RT can be measured, where the second task is thought to require all the mental operations of the first, plus an additional inserted operation. The difference between mean RTs in the two tasks is interpreted as an estimate of the duration of the inserted stage, as shown in Fig. 1. This interpretation depends on the validity of both the stage theory and an *assumption of pure insertion* which states that changing from task 1 to Task 2 merely

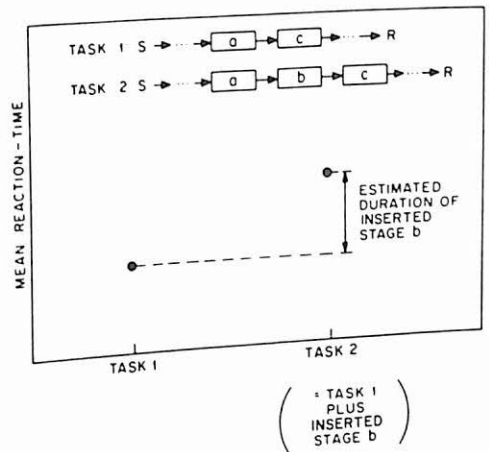


Figure 1
Donders' subtraction method. Hypothetical stages between stimulus (S) and response (R) are represented by a, b, and c.

inserts a new processing stage without altering the others.

For example, Wundt (1880, pp. 247–260) developed an application in which RTs were measured when a subject had to respond after he had identified a stimulus, and also when he had to respond after merely detecting its presence. The difference was used as an estimate of the identification time. In this instance the stages shown in Fig. 1 might be (a) stimulus detection, (b) stimulus identification, and (c) response organization. In an earlier application, Donders (1868) had compared mean RTs in a simple-reaction task (one stimulus and response) and a choice-reaction task (multiple stimuli and responses); he regarded the difference as the duration of the stages of stimulus discrimination and response selection.

This kind of enterprise occupied many psychologists during the last quarter of the 19th century. Much of their work was summarized by J. Jas-

trow (1890) in a popular treatise on *The Time Relations of Mental Phenomena*.

Around the turn of the century the subtraction method became the subject of criticism, for two main reasons. First, the differences in mean RT that were observed in some applications varied excessively from subject to subject, and from laboratory to laboratory. In retrospect, this seems to have been caused by the use of tasks and instructions that left the subject's choice of "processing strategy" relatively uncontrolled.² Second, introspective reports put into question the assumption of pure insertion, by suggesting that when the task was changed to insert a stage, other stages might also be altered. (For example, it was felt that changes in stimulus-processing requirements might also alter a response-organization stage.) If so, the difference between RTs could not be identified as the duration of the inserted stage. Because of these difficulties, Külpe, among others, urged caution in the interpretation of results from the subtraction method (1895, Secs. 69, 70). But it appears that no tests other than introspection were proposed for distinguishing valid from invalid applications of the method.

A stronger stand was taken in later secondary sources. For example, in a section on the "discarding of the subtraction method" in his *Experimental Psychology* (1938, p. 309), R. S. Wood-

worth said "[Since] we cannot break up the reaction into successive acts and obtain the time of each act, of what use is the reaction-time?" And, more recently, D. M. Johnson said in his *Psychology of Thought and Judgment* (1955, p. 5), "the reaction-time experiment suggests a method for the analysis of mental processes which turned out to be unworkable."

Nevertheless, the attempt to analyze RT into components goes on, and there has been a substantial revival in the last few years in the use of RT as a tool for the study of mental processes ranging from perceptual coding to mental arithmetic and problem-solving.³ The work on memory retrieval described here is part of this revival, and is based heavily on Donders' stage theory. Modern styles of experimentation and data analysis lead to applications of the stage theory that seem to withstand the early criticisms, and to tests of validity other than introspection.

I shall describe experiments on retrieval from memory that have led to the discovery of some relatively simple search processes. My aim is to convey the general outline rather than the details of this work, so the picture I paint will be somewhat simplified; there will be little discussion of alternative explanations that have been considered and rejected. Such discussions can be found in Sternberg (1966, 1967a, b, and 1969).

The purpose of most of these experiments has been to study the ways in which information is retrieved from memory when learning and retention are essentially perfect. The method is to present a list of items

² For example, Cattell (1886, p. 377) reported that "I have not been able myself to get results by [Wundt's] method. I apparently either distinguished the impression and made the motion simultaneously, or if I tried to avoid this by waiting until I had formed a distinct impression before I began to make the motion, I added to the simple reaction not only a perception, but a volition."

³ See, e.g., Egeth, 1966; Hochberg, 1968; Nickerson, 1967; Posner & Mitchell, 1967; Restle & Davis, 1962; Smith, 1967; Suppes & Groen, 1966.

for memorization that is short enough to be within the immediate memory span. The subject is then asked a question about the memorized list; he answers as quickly as he can, and his delay in responding is measured. By examining the pattern of his RTs, while varying such factors as the number of items in the list and the kind of question asked, one can make inferences about the underlying retrieval processes. Since the aim has been to understand error-free performance, conditions and payoffs are arranged so that in most experiments the responses are almost always correct.

2. Judging Presence Versus Absence in a Memorized List

The flavor of this approach will become clearer as we consider a particular experiment. Figure 2 shows the paradigm of an *item-recognition task*. The stimulus ensemble consists of all potential test stimuli. From among these, a set of s elements is

selected arbitrarily and is defined as the *positive set*; these items are presented as a list for the subject to memorize. The remaining items are called the *negative set*. When a test stimulus is presented, the subject must decide whether it is a member of the positive set. If it is, he makes a *positive response* (e.g., saying "yes" or operating a particular lever). If not, he makes a *negative response*. The measured RT (sometimes referred to as *response latency*) is the time from test-stimulus onset to response.

Within the item-recognition paradigm, different procedures can be used. One of them, shown at the top of Fig. 3, is the *varied-set procedure*. Here, the subject must memorize a different positive set on each trial. In one experiment (Exp. 1), for example, the stimulus ensemble consisted of the ten digits. On each trial a new positive set, ranging randomly over trials from one to six different digits, was presented sequentially at a rate of 1.2 seconds per digit. Two seconds after the last digit in the set was displayed, a warning signal appeared, followed

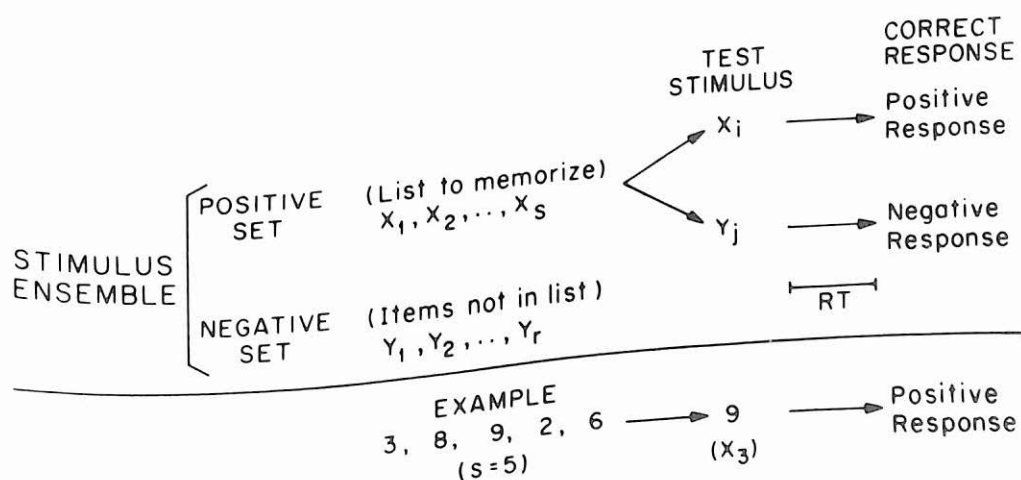
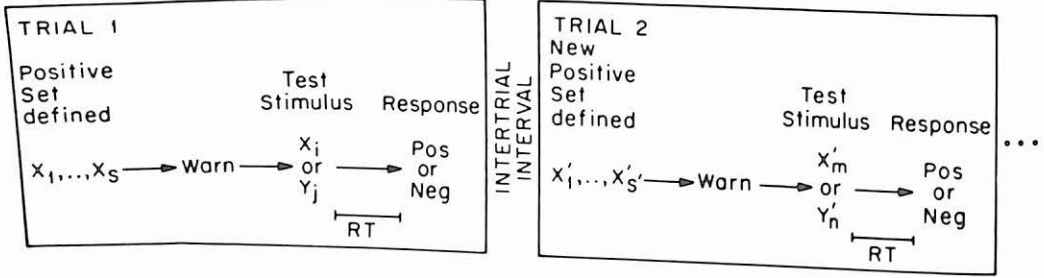


Figure 2

Paradigm of item-recognition task (Exps. 1-5).

VARIED-SET PROCEDURE



FIXED-SET PROCEDURE

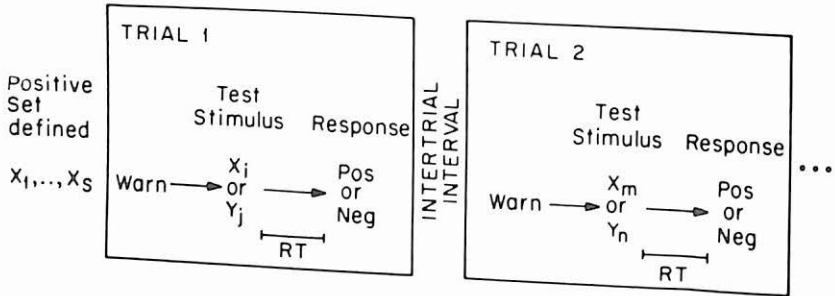


Figure 3

Varied-set and fixed-set procedures in item-recognition. A Y represents an item in the negative set. Primes are used in representing trial 2 of the varied-set procedure to show that both the items in the positive set ($X_1, \dots, X_s$) and its size (s) may change from trial to trial.

by a visually-presented test digit. The subject pulled one lever, making a positive response, if the test stimulus was contained in the memorized list. He pulled the other lever, making a negative response, if it was not. After responding to the test stimulus the subject recalled the list. This forced him to retain the items in the presented order, and prevented him from working with the negative set rather than the positive. Regardless of the size of the positive set, the two responses were required equally often. As in the other experiments I shall describe, subjects were relatively unpracticed. The error rate in this kind of experiment can be held to 1 or 2 percent by paying subjects in such a

way as to penalize errors heavily while rewarding speed.

Averaged data from eight subjects are shown in Fig. 4. Mean RT is plotted as a function of the number of symbols in memory—that is, the number of digits in the positive set that the subject committed to memory at the start of the trial.

These data are typical for item recognition experiments. They show, first, a linear relation between mean RT and the size of the positive set. Second, the latencies of positive and negative responses increase at approximately the same rate. The slope of the line fitted to the means is 38 msec. per item in memory; its zero-intercept is about 400 msec. (It hap-

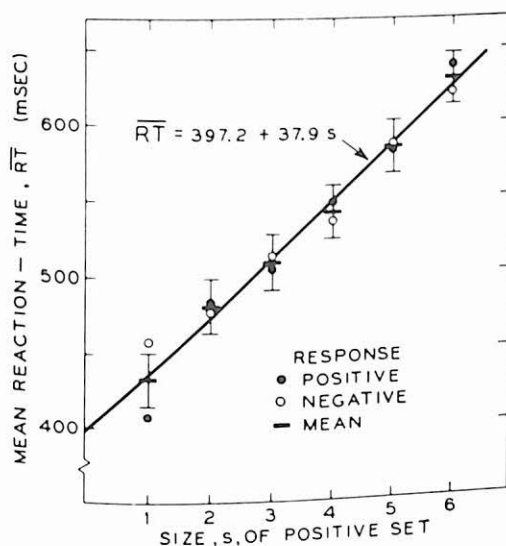


Figure 4

Results of Exp. 1: Item-recognition with varied-set procedure. Mean latencies of correct positive and negative responses, and their mean, as functions of size of positive set. Averaged data from eight subjects, with estimates of $\pm\sigma$ about means, and line fitted by least squares to means.

pens to be true in these data that latencies of positive and negative responses have approximately the same values: the two latency functions have not only the same slope but also the same zero-intercept. This is not a general finding, but results from the particular conditions in this experiment. By varying the relative frequency with which positive and negative responses are required, for example, one can vary the relation between their latencies. But as relative frequency is varied the slopes of the two latency functions remain equal and unchanged.) Before considering the interpretation of these findings, we turn to some general matters regarding search processes.

3. Two Types of Serial Search

Let *serial search* (or *scanning*) be a process in which each of a set of items is compared one at a time, and no more than once, to an item. Linear RT-functions, as in Fig. 4, suggest that subjects in the item-recognition task use a serial search process whose mean duration increases by one unit for each additional comparison. The purpose of the search is to determine whether an agreement (or *match*) exists between the test item and any of the items in the memorized set. Two types of serial search that might serve this purpose need to be considered. In *self-terminating serial search*, the test stimulus is compared successively to one item in memory after another, either until a match occurs (leading to a positive response), or until all comparisons have been completed without a match (leading to a negative response). In *exhaustive serial search*, the test stimulus is compared successively to *all* the memorized items. Only then is a response made—positive if a match has occurred, and negative otherwise. A self-terminating search might require a separate test, after each comparison, to ascertain whether a match had occurred, rather than only one such test after the entire series. On the other hand, an exhaustive search must involve more comparisons, on the average, than a search that terminates when a match occurs.

Suppose that the average time from the beginning of one comparison to the beginning of the next is the same for each comparison in the series, and is not influenced by the number of comparisons to be made. Then the durations of both kinds of search will increase linearly with the number of

memorized items (*list length*). There are, however, important differences. In an exhaustive search the test stimulus is compared to all items in memory before each positive response as well as before a negative response. Hence, the rate at which RT increases with list length—the slope of the RT-function—is the same for positive and negative responses. In contrast, self-terminating search stops in the middle of the list, on the average, before positive responses, but continues through the entire list before negatives. The result is that as list length is increased, the latency of positive responses increases at half the rate of the increase for negatives. This difference between the two kinds of search is illustrated on the left side of Fig. 5.

A second difference between the two types of search, illustrated on the right side of Fig. 5, is in the

serial-position functions for positive responses. In a simple exhaustive search neither the order of search nor the position of the matching item in the list should have any effect on the RT, since all items are compared. A self-terminating search that occurred in a random order, or started at a random point, also would produce flat serial-position curves. But if a self-terminating search started consistently with the first item, and proceeded serially, then the serial-position curves would increase linearly. (If, in addition, list length influenced *only* the search process, then the curves for different list lengths would be superimposed: for example, the time to arrive at the second item in a memorized list would be independent of the length of the list.) Increasing serial-position functions are therefore sufficient (but not necessary) evidence for

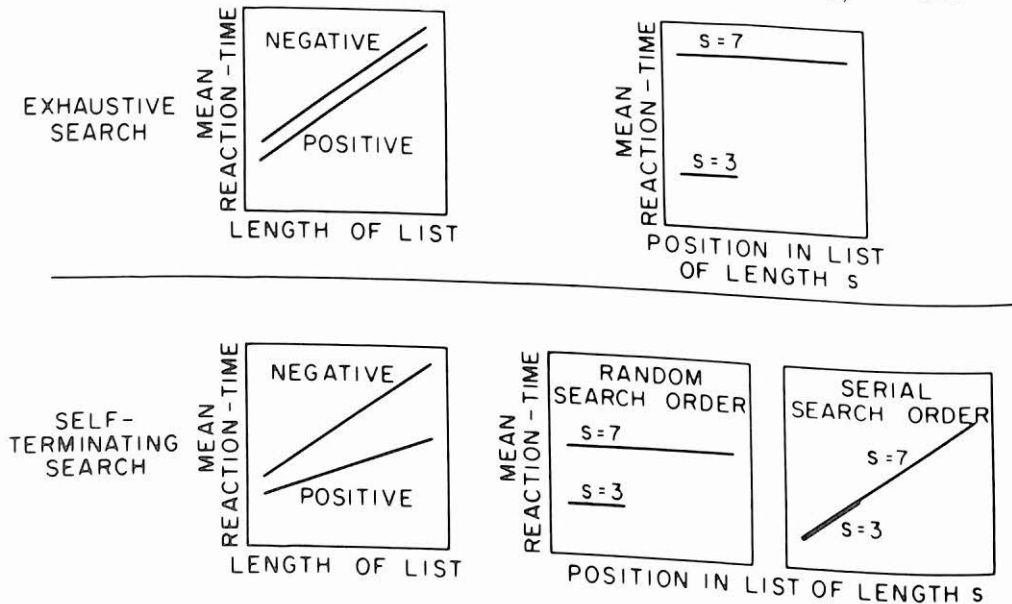


Figure 5

Some properties of exhaustive (top) and self-terminating (bottom) serial search. Left: Theoretical RT-functions (mean latencies of positive and negative responses as functions of length of list). Right: Theoretical serial-position functions (mean latency of positive responses as a function of serial position of test item in a list of given length).

inferring that a search process is self-terminating.

4. High-Speed Exhaustive Scanning

The serial-position curves actually observed in the item-recognition experiment described in Sec. 2 were relatively flat.⁴ Together with this finding, the linearity of the latency functions and the equality of their slopes for positive and negative responses indicate an exhaustive search. The data show also that memory-scanning can proceed at a remarkably high rate. The slope of the mean RT-function, which is an estimate of the time per comparison, was 38 msec., indicating an average scanning rate between 25 and 30 digits per second.

Perhaps because of its high speed, the scanning process seems not to have any obvious correlate in conscious experience. Subjects generally say either that they engage in a self-terminating search, or that they know immediately, with no search at all, whether the test stimulus is contained in the memorized list.

Is high-speed scanning used only when a list has just been memorized and is therefore relatively unfamiliar? The results discussed so far

(Fig. 4) are from the varied-set procedure (Fig. 3), in which the subject must memorize a new positive set on each trial, and is tested only three seconds after its presentation. How is the retrieval process changed when a person is highly familiar with a particular positive set and has had a great deal of practice retrieving information from it? At the bottom of Fig. 3 is shown the *fixed-set procedure* in the item-recognition paradigm, in which the same positive set is used for a long series of trials. For example, in one experiment (Exp. 2) subjects had 60 practice trials and 120 test trials for each positive set. On the average test trial, a subject had been working with the same positive set for ten minutes, rather than three seconds. The sets were sufficiently well learned that subjects could recall them several days later. Sets of one, two, and four digits were used. There were six subjects.

Results are shown in Fig. 6, and are essentially identical to those from the varied-set procedure. The RT data are linear, the slopes for positive and negative responses are equal, and the average slope is 38 msec. per digit. The small difference between the zero-intercepts in the two experiments is not statistically significant. The remarkable similarity of results from the two procedures indicates that the same retrieval process was used for both the unfamiliar and the well-learned lists.

5. Active and Inactive Memory

Evidence has accumulated, particularly during the past decade, that there are at least two systems or states of memory for encoded verbal items (e.g., Broadbent, 1958; Waugh & Norman, 1965; Glanzer & Cunitz, 1966; Atkin-

⁴ Several investigators have, however, reported marked recency effects in item-recognition tasks: RTs were shorter for test stimuli later in the list (Corballis, 1967; Morin, DeRosa & Stultz, 1967; Morin, DeRosa & Stultz, 1967; Morin, DeRosa & Stultz, 1967). Without embellishment a theory of exhaustive scanning cannot, of course, handle such findings. The salient procedural characteristics of experiments that produce such recency effects seem to be a fast rate of list presentation and a short interval (less than 1 sec.) between the last item in the list and the test item. Findings of Posner et al. (1969) indicate that in this range the time interval between successive stimuli may critically influence the nature and duration of comparison operations.

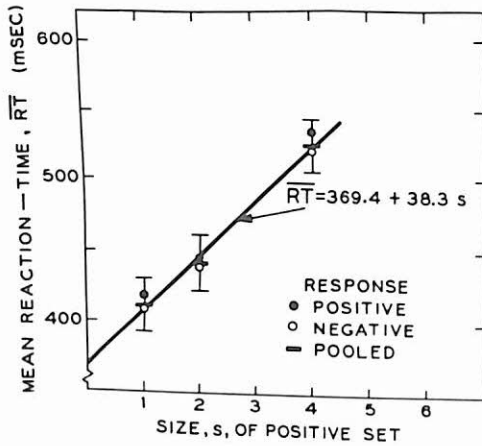


Figure 6

Results of Exp. 2: Item-recognition with fixed-set procedure. Mean latencies of correct positive, negative, and pooled responses as functions of size of positive set. Averaged data from six subjects, with estimates of $\pm\sigma$ about pooled means, and line fitted by least squares to those means. For each set size positive responses were required on 27% of the trials.

son & Shiffrin, 1968). The picture that is emerging is roughly as follows: The *long-term store*, or *inactive memory*, is relatively permanent and of large capacity. It receives information from the *short-term store*, a temporary *active memory*⁵ of small capacity from which information is rapidly lost unless an active retention process is operating. In the long-term store, the coding of verbal items includes semantic attributes; in the short-term store, however, such items are coded primarily as acoustic or articulatory

⁵ An alternative term is "working memory," used by Newell & Simon, 1963, to refer to the arithmetic unit of a general-purpose computer. See also "Active verbal memory," Ch. 9 in Neisser, 1967, and "Operational memory," Sec. 4 in Posner, 1967.

representations of their spoken names, even when they have been presented visually (see Sperling, 1960; Conrad, 1964; Baddeley, 1966; Wickelgren, 1969). The active process that regenerates the rapidly-decaying traces of a list of items in the short-term store is *rehearsal*, the overtly- or silently-spoken cyclic serial recall of stored items (see Sanders, 1961; Sperling, 1963; Posner & Rossman, 1965; Cohen & Johansson, 1967; Crowder, 1967; Atkinson & Shiffrin, 1968). Rehearsal, which also causes information in the short-term store to be entered in the long-term store, has an approximate maximum rate of from three to seven items per second (Landauer, 1962).

Whereas in the varied-set procedure of Exp. 1 the positive set must have been stored in active memory only, it is reasonable to believe that the set had entered the long-term store in the fixed-set procedure of Exp. 2. However, the similarity of results from the two procedures suggests that the same memory system was being scanned: that is, when information in inactive memory has to be used, it may be entered also in active memory (where it is maintained by rehearsal) and thus become more readily available. An experiment that tests this conjecture is described below (Exp. 5).

It appears, then, that the memory of the positive sets in both tasks is maintained by a serial rehearsal process; supporting this notion, subjects reported silent rehearsal of the sets in both experiments. But the estimated rates of high-speed scanning and the fastest silent speech differ by a factor of at least four. Rehearsal is far too slow to be identical to the scanning process. Instead, it should be thought of as a separate process whose only function in these tasks is

to maintain the memory that is to be scanned.⁶

6. Encoding of the Test Stimulus

In the scanning process inferred from these experiments, some internal representation of the test stimulus is compared to internal representations of the items in the positive set. What is the *nature* of the representations that can be compared at such high speed? Another way to phrase the question is to ask how much processing of the test stimulus occurs before it is compared to the memorized items.

Various considerations lead one to expect a good deal of preprocessing. For example, the idea that items held in active memory are retained as acoustic or articulatory representations of their spoken names introduces the possibility that the test stimulus is processed to the point of naming, and that the name of the test stimulus is compared to the names of the items in the positive set. But of the items in the positive set. But two points should be kept in mind regarding this possibility. First, it would require that stored names could be scanned much faster than they could be covertly articulated, since the scanning rate is about four times as fast as people can say names of digits to themselves. Second, unlike other

forms of preprocessing, such as image-sharpening or feature-extraction, preprocessing a character to the point of identification or naming would itself require the retrieval of information from memory—that information which relates the character to its name.

In one experiment bearing on this question (Exp. 3), I degraded the test stimulus by superimposing a pattern that had been adjusted to increase the RT without substantially altering the error rate. I then examined the effect of stimulus quality on the function that relates mean RT and the size of the positive set. It is shown below that this effect would depend on the nature of the internal representation of the test-stimulus.

Figure 7 shows idealized data from a scanning experiment. The zero-intercept corresponds to the total duration of all processes.

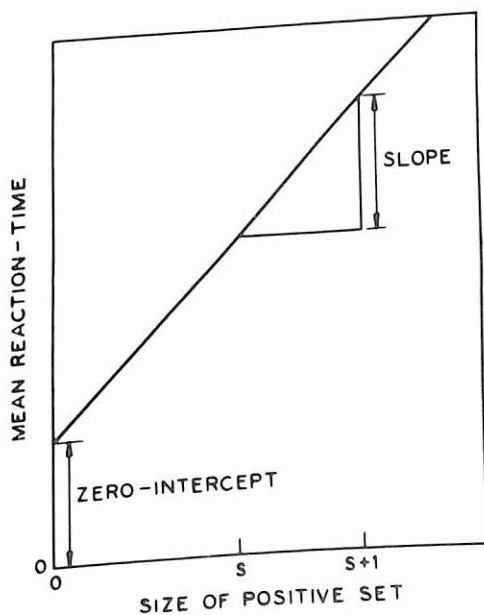


Figure 7

Idealization of mean RT-function from an item-recognition task.

⁶ It is sometimes thought that the six or seven objects in the "span of apprehension" are immediately and simultaneously available, being contained in the "psychological present." And the information in active memory has occasionally been identified with this momentary capacity of consciousness (e.g., Miller, 1962, pp. 47-49; Waugh & Norman, 1965). The finding that one must scan one's active memory to ascertain its contents, rather than having immediate access to them, reveals a possible flaw in this argument.

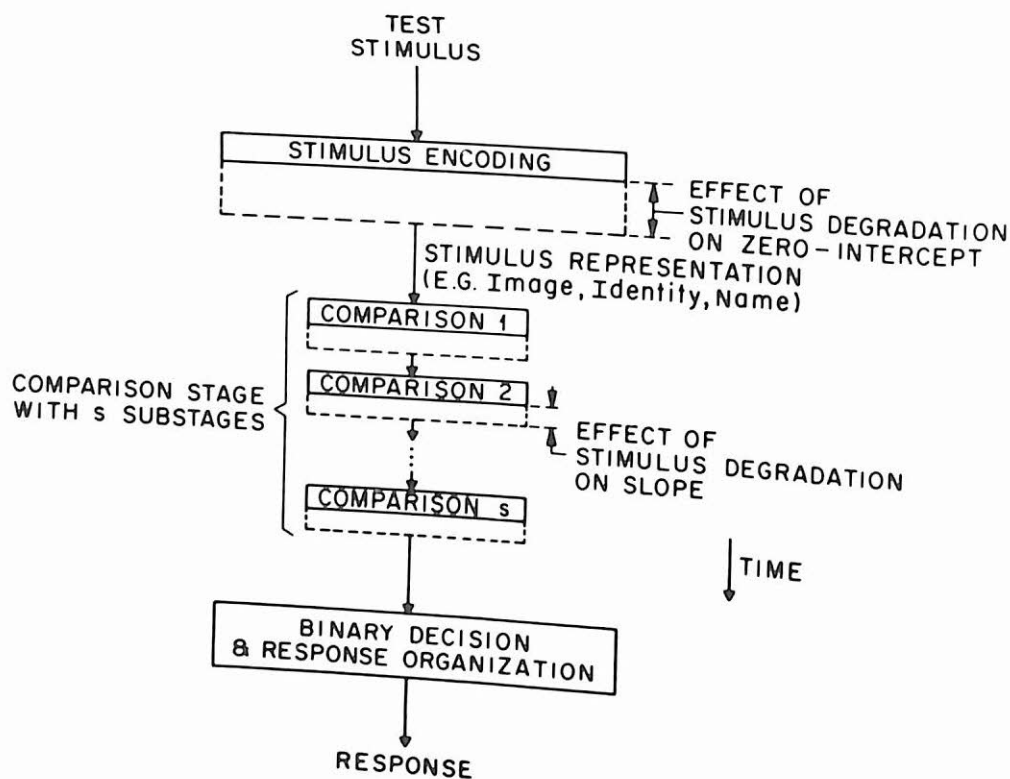


Figure 8
Some hypothetical stages and substages in item-recognition, and two possible effects of test-stimulus quality on stage and substage durations. Height of box represents mean duration of that stage or substage.

that occur just once, regardless of the size of the positive set—such as the encoding of the test stimulus to form its representation, and the organization and execution of the motor response. The slope, on the other hand, measures the duration of processes that occur once for each member of the positive set—the comparison operation, and the time to switch from one item to the next.⁷ Figure 8 shows

a flow diagram of some hypothetical stages between test stimulus and response. The height of a box represents the mean duration of that stage. An effect of stimulus degradation on the *stimulus-encoding stage*, which generates the stimulus representation, would increase the zero-intercept of the RT-function. An effect on the

⁷ This analysis assumes that the mean durations of comparisons leading to matches and to mismatches are equal. Without this assumption all the statements here (and elsewhere in the paper) are correct, except that the slope of the RT-function measures the

mean duration of only those comparisons that lead to mismatches, together with the time to switch from one comparison to the next. Any difference between durations of the two kinds of comparison would contribute to a difference between zero intercepts of the latency functions for positive and negative responses.

serial-comparison stage would increase the slope, since a time increment would be added for each item compared.

Consider two extreme possibilities: First, suppose that the encoding stage did nothing other than transmit an unprocessed image, or direct copy, of the test stimulus. Then degradation could influence only the comparison operation, which occurs once for each member of the positive set; only the slope of the RT-function would change, as in Panel A of Fig. 9. At the other extreme, suppose that the representation produced by the encoding stage was the *name* of the test stimulus. The input to the serial-comparison stage would be the same, whether or not the test stimulus had been degraded by a superimposed visual pattern; hence degradation could not influence this stage. (For the serial-comparison stage to be influenced by visual degradation, its input would have to be visual, in the sense of embodying details of the physical stimulus pattern that are not present in the mere name of the stimulus.) Only the encoding stage, then, could be influenced by degradation; and since encoding takes place just once, only the zero-intercept of the RT-function

would change, as in Panel B of Fig. 9. (The absence of a change in slope, however, does not necessarily imply a non-visual stimulus-representation; the representation could be visual, but highly processed.)

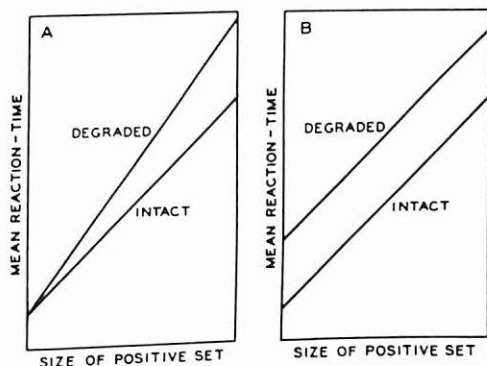


Figure 9

Two possibilities for the effect of test-stimulus quality on the RT-function. A: Quality influences comparison stage only. B: Quality influences encoding stage only.

In Exp. 3 each of twelve subjects had positive sets of one, two, and four digits, with test stimuli *intact* in some blocks of trials, and in others *degraded* by a superimposed checkerboard pattern. Intact and degraded numerals are shown in Fig. 10.

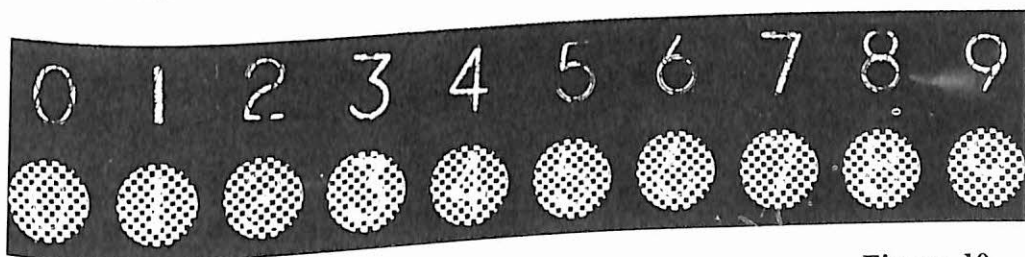


Figure 10

Photographs of intact and degraded numerals used in Exp. 3. Numerals were about 0.6 in. high and were viewed from a distance of about 29 in. Degraded numerals were somewhat more discriminable than they appear in the black-and-white photograph, possibly because of a slight color difference between numerals and checkerboard.

The fixed-set procedure was used. Results for the two sessions are shown separately in Fig. 11. Consider first the data from the second session, on the right-hand side of the figure. Latencies of positive and negative responses have been averaged together. The functions for degraded and intact stimuli are almost parallel, but there is a large effect on the zero-intercept, closely approximating the pattern shown in Panel B of Fig. 9. This indicates that degradation had a large influence on the stimulus-encoding stage, and that the representation generated was sufficiently processed that the serial-comparison stage could proceed as rapidly with degraded as with

intact stimuli. The stimulus representation was either nonvisual or, if visual, sufficiently refined in the second session to eliminate any effect of degradation.

The data from this session are an instance of the *additivity* of two effects on RT. There is no interaction between the effect of set size and the effect of stimulus quality; instead, the effect of each of these factors on mean RT is independent of the level of the other. Such additivity supports the theory of a sequence of stages, one stage influenced by stimulus quality and the other by set size (see Sec. 7).

Now let us consider the data from the first session, shown on the left-hand side of Fig. 11. Here, where subjects have not yet had much practice with the superimposed checkerboard, there is a 20% increase in the slope of the RT-function, as well as an increase in its zero-intercept. This pattern agrees with neither of the pure cases of Fig. 9. Stimulus quality apparently *can* influence the duration of comparison operations; hence, the output of the encoding stage must be sensitive to degradation. Findings from the two sessions imply, then, that although the stimulus representation is highly processed, it embodies physical attributes of the test stimulus, rather than being a name or identity. That is, the test-stimulus representation is visual. The memory representations of the positive set that are used in the serial-comparison stage must therefore also be visual, to make comparison possible. Hence, although items in the positive set appear to be represented as covertly-spoken names in the course of their rehearsal, this is not the only form in which they are available.

What changed between the first and second sessions so as to virtually elim-

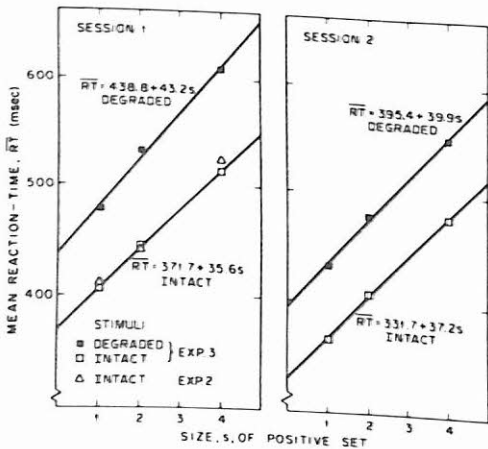


Figure 11

Results of Exp. 3: Effect of stimulus quality on item-recognition. Mean RT, based on pooled data from positive and negative responses, as a function of size of positive set for intact and degraded test stimuli. Left-hand and right-hand panels show data from Sessions 1 and 2, respectively. Averaged data from 12 subjects, with lines fitted by least squares. In all conditions positive responses were required on 27% of the trials. Triangles show results from Exp. 2 (Fig. 6), which was similar.

inate the influence of stimulus quality on the slope of the RT-function? Since the scanning rate with intact stimuli and the effect of degradation on the zero-intercept are approximately the same in the two sessions, it seems unlikely that the type of representation changed. For the present, my interpretation is that the encoding stage became more efficient at removing the effects of the fixed degrading pattern.

Additional support for the idea that the memory representations scanned in the item-recognition task have sensory characteristics, rather than being completely abstracted from the physical stimuli, comes from two other studies. In the first, Chase and Calfee (1969) created four different conditions in the varied-set procedure by presenting both the positive set and the test stimulus either visually or aurally. When the set and test item were presented in different modalities, the slope of the RT-function increased by about 30%, indicating a slower scanning rate. If abstract representations were being compared in the same-modality conditions, then the change to different-modality conditions should have altered only the zero-intercept, as in Fig. 9B. In the second study, Posner et al. (1969) concluded that when a single letter is presented *aurally* for memorization, the decision whether a *visual* test-letter is the "same" is facilitated by the internal generation of a visual representation of the memorized letter, which obviates the need to identify the test letter. Still further evidence will be discussed below (Exp. 4).

7. A Test of the Stage Theory

The work described above is grounded on Donders' stage theory.

That is, as in his subtraction method, the effects on mean RT of changes in experimental conditions (factors) have been attributed to the selective effects of these factors on hypothetical processing stages between stimulus and response. How can we ensure that such inferences are not open to the classical criticism of the subtraction method, that even if information processing is organized in functionally different stages, factor effects may not be selective? One answer, of course, is that the test of a method's applicability is whether it produces results that fit together and make sense. But there are two other arguments as well.

The first stems from replacement of the assumption of pure insertion by a weaker and more plausible *assumption of selective influence*. Instead of requiring that a change in the task insert or delete an entire processing stage without altering others, the weaker assumption requires only that it influence the *duration* of some stage without altering others. One example is illustrated in Fig. 12. To estimate the comparison time by the subtraction method, one would have studied Task 2, in which the positive set has one member, and compared it to Task 1.

Task 1 would have been constructed to measure the zero-intercept directly, by deleting the entire comparison stage. But I suspect that there is no appropriate Task 1, in which deletion of all comparisons would leave the other stages of processing invariant. In this instance, then, the assumption of pure insertion is probably invalid. This is why the important RT-differences in the experiments described above were those between Tasks 2 and 3, 3 and 4, and so on, whose interpretation required only that the comparison stage be selectively in-

fluenced by set size. Similarly, in studying the preprocessing of the stimulus, instead of entirely eliminating the need to discriminate the stimulus (in an effort to *delete* the hypothetical encoding stage) I examined the effects of making its discrimination more or less difficult, thereby varying the amount of work the stage had to accomplish. Of course, one result of using a factor that influences but does not insert a stage is that we have no estimate of the stage's total duration. But that seems to be of less interest than whether there is such a stage, what influences it, what it accomplishes, and what its relation is to other stages.⁸

In a given experimental situation the validity of even the weaker assumption of selective influence must be checked, however. We can distinguish those situations where one of the assumptions—*influence* or *insertion*—holds, by testing the additivity of the effects of two or more factors on mean RT (Sternberg, 1969). It is this test that provides the second and most telling way of dealing with the classical criticism. Consider a pair of hypothetical stages and a pair of experimental factors, with each factor inserting or selectively influencing one of the stages. Because stage durations are additive (by definition), the changes in mean RT produced by such factors should be independent and additive. That is, the effect of one factor will be the same at all levels of the other, when the response is meas-

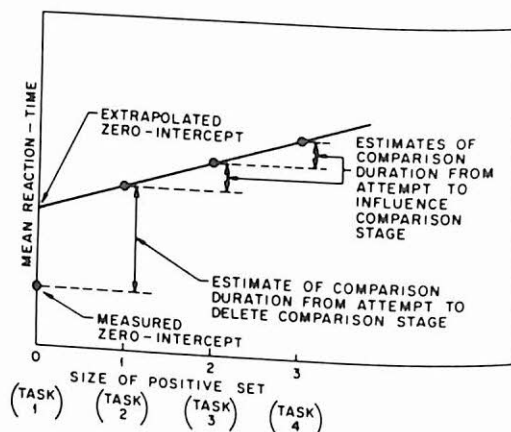


Figure 12
Example of error from hypothetical attempt to estimate comparison time by deleting the comparison stage altogether, as in the subtraction method, and to use a measured zero intercept. Attempt fails because *deletion* of comparison stage changes the demands placed on other stages, whereas *variation* of the number of comparisons, s , ($s \geq 1$) does not.

ured on a scale of time or its (arithmetic) mean.⁹

In experiments with the fixed-set procedure I have examined four factors, which are listed above the broken line in Fig. 13. The additivity of five of the six possible factor pairs has been tested and confirmed (1 and 2, and 3, 2 and 4, and 3 and 4). These instances of additivity support the assumption that the factors selectively influence different stages of processing and, *a fortiori*, confirm the existence of such stages. Another instance of additivity, and the one on which inferences about the structure of the

⁸ This alternative was preferred by Cattell, 1886, who argued (p. 378) "I do not think it is possible to add a perception to the reaction without also adding a will-act. We can however change the nature of the perception without altering the will-time, and thus investigate with considerable thoroughness the length of the perception time." But he suggested no way to test these assertions.

⁹ Discussions of various other aspects and modern versions of the subtraction method, including considerations of validity, may be found in Hohle, 1967; McGill & Gibbon, 1965; McMahon, 1963; Smith, 1968; Sternberg, 1964; Sternberg, 1969; and Taylor, 1966.

the effect of the number of letters in the set.

Another possible alternative to serial search is an "associative" process. Consider the case in which positive items are distinguished by membership in a well-learned category. (For example, the positive set might contain digits only, and the negative set, letters.) To each member of a category is associatively linked its category label, and the binary choice depends on which label is elicited by the test stimulus. The speed of such a process might be independent of the sizes of positive and negative sets (although it might depend on various attributes of the categories that contained them, including *their* sizes; see Landauer & Freedman, 1968). On the other hand, the high speed of scanning might make it more efficient than an associative process when one of the sets is small. In short, there may be alternative mechanisms for the same task, and which one is used may depend, in part, on which one is more efficient. If this is the case it is a great advantage to understand at least one of the alternatives in some detail.

9. Retrieval of Nonsymbolic Versus Symbolic Information

Other questions about the generality of the scanning process are raised by its high speed, which precludes its being identified with the subvocalization of numeral names, and also by the influence of stimulus-quality on the scanning rate in Exp. 3 (Sec. 6), which indicates that the stimulus representation is not the name or identity of the numeral. The fact that numerals are patterns with extremely well-learned names may therefore be irrelevant to the scanning process. Of course, numerals have other spe-

cial properties: they are highly familiar, they are symbols, they represent numerical quantities, and people are practiced at manipulating the numbers they represent. A. M. Treisman and I recently tested the importance of these properties for memory retrieval, using two different ensembles, one of nonsense forms, and one of the other photographs of faces (Exp. 4). To our subjects, both ensembles were unfamiliar, nonsymbolic, unordered, and without well-learned names. We used the fixed-set procedure with sets of size 1 to 4, but found it necessary to display the positive set before each trial in order to help the subjects, who were inexperienced, to maintain it in active memory.

RT data, shown in Fig. 14, are qualitatively the same as those for digit sets. They show linearity, suggesting a serial process, and equality of slopes for positive and negative responses, indicating exhaustiveness of search. The main difference is in the scanning rate, which seems to depend to some extent on the nature of the stimuli. Even for faces, however, the estimated rate is high—about 18 faces per second. These findings indicate that high-speed exhaustive scanning does not depend on the special properties of numerals mentioned above. They also add further support to the conclusion that the test-stimulus representation in the case of numerals is not the name of the numeral, but is some sort of visual representation.

10. Retrieval from Inactive Versus Active Memory

A further question about the generality of the high-speed scanning process is raised by the conjecture (Sec. 5) that it occurs only when in-

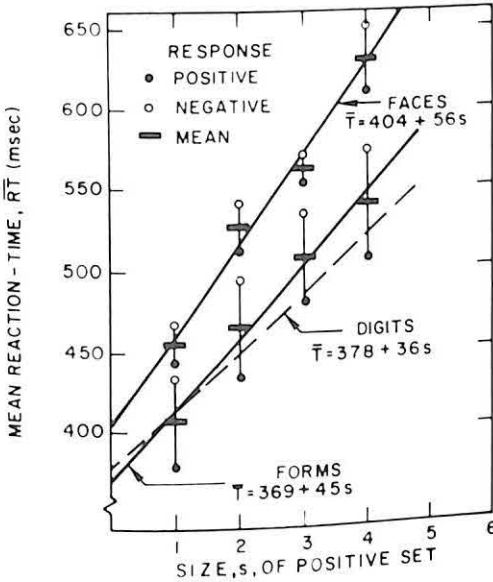


Figure 14

Results of Exp. 4: Item-recognition with nonsense forms and photographs of faces. Mean latencies of correct positive and negative responses, and their mean, as a function of size of positive set for the two stimulus ensembles. Averaged data from eight subjects for each ensemble, with lines fitted by least squares to means. Broken line was fitted to data from a similar experiment with an ensemble of numerals.

formation is being held in active memory. The similarity of results from the varied-set and fixed-set procedures led to the idea that even when a list is contained in long-term memory, it is transferred into active memory and maintained there by rehearsal in order to be used in the item-recognition task. If that is so, one would expect some change in the process if one prevented the relevant list from being rehearsed (for example, by occupying the active memory with other material). This kind of procedure moves us closer to studying the differences between retrieval from the short-term (active) and long-term (inactive) stores, and thereby understanding the latter by using the former as a baseline.

The procedure in a small preliminary experiment (Exp. 5) is shown in Fig. 15. At the start of a series of trials the subject memorized a list of 1, 3, or 5 digits, which defined the positive set for the entire series. On each trial a new list of seven letters was presented sequentially, at a rate of two letters per second. A short time after the last letter, there was a brief warning signal, and then one

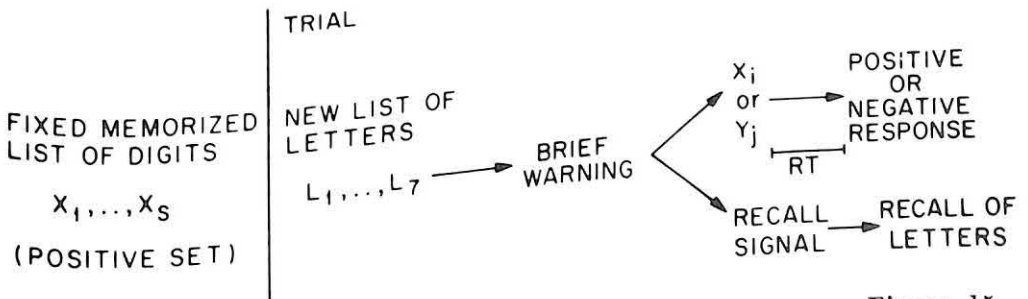


Figure 15

Paradigm of Exp. 5: Item-recognition from active and inactive memory. Only the inactive-memory condition is shown. In the active-memory condition, also involving a fixed-set procedure (Fig. 3), no letters were presented.

of two things could happen. On a random third of the trials the subject saw a recall signal, and attempted to recall the seven letters. These trials were used in order to encourage the subject to attend to the letters and retain them in memory until the test event. (Observing and retaining the list of letters was intended to occupy his active memory on all trials and prevent him from rehearsing the positive set.) On the remaining trials the subject saw a test digit. He was required to make a positive or negative response, based on the previously memorized digit set, as quickly as possible consistent with accuracy. This is a difficult task, and required a session of practice for smooth performance. In the series of control trials, which alternated with series of experimental trials, no lists of letters were presented.

Data averaged over the four subjects in this preliminary experiment are shown in Fig. 16. The lower set of points represents performance in the control condition, which was similar in procedure to Exp. 2. Mainly because of one exceptional subject, the fitted line is somewhat steeper than usual, with a slope of 57 msec. per digit. Otherwise the data are typical. In the experimental condition the fitted line is about twice as steep as in the control condition, with a slope of 105 msec. per digit. Again, the latencies of positive and negative responses grow at equal rates as set size is increased. The zero-intercepts in the experimental and control conditions differ also, by over 100 msec.

Evidently, the retrieval process is radically altered, with the effective scanning rate halved, when the information to be retrieved is not being rehearsed and is therefore not in active memory. Current notions about the functions of rehearsal include

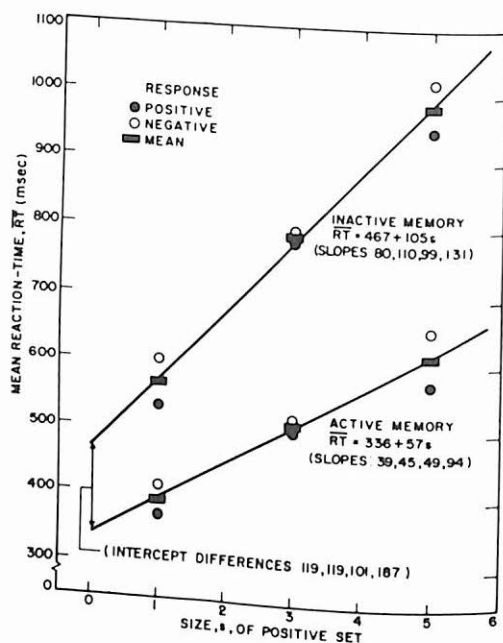


Figure 16 Results of Exp. 5: Item-recognition from active and inactive memory. Mean latencies of correct positive and negative responses, and their mean, as functions of size of positive set, in conditions of active and inactive memory. Averaged data from four subjects, with lines fitted by least squares to means. Intercept differences and slopes for the four subjects are listed, the order of subjects being the same in each list.

maintenance of short-term memory, and transfer of information into long-term memory (see Sec. 5). The results of Exp. 5 suggest a third role—that of making information already stored in long-term memory more rapidly accessible.

At this point there is little basis for selecting among potential explanations for the data from the experimental condition, but experiments are under way that may help to do so. The explanation that I favor, shown in Fig. 17, is the one that

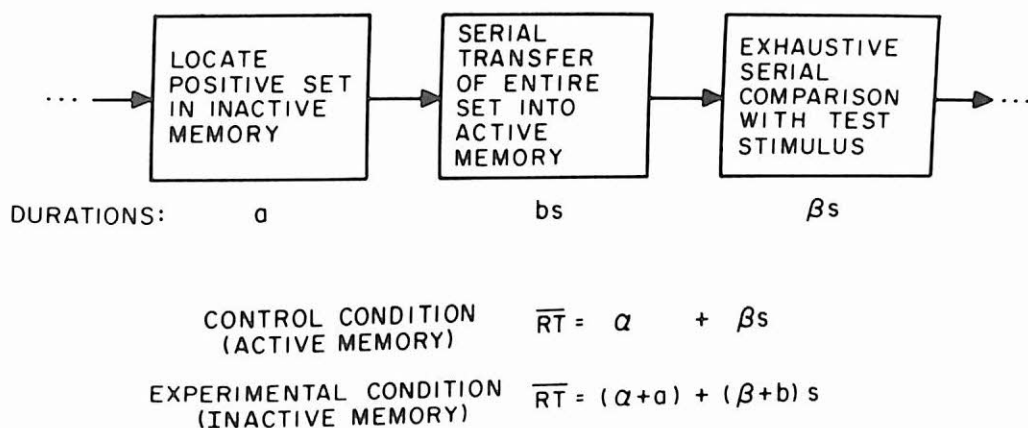


Figure 17

One explanation of results of Exp. 5. Left and middle boxes represent hypothetical stages that might be interpolated in the inactive-memory condition. Also shown are hypothesized durations of these two stages and the comparison stage, and resulting theoretical RT-functions in which α represents the zero-intercept of the RT-function in the active-memory condition.

makes plausible two striking aspects of the data: despite the large effect of condition, the linearity of the RT-function, and the equality of slopes for positive and negative responses, are both preserved. The first two boxes in the figure represent hypothetical stages that might be present in the experimental condition but not in the control. One might be searching for the positive set in inactive memory. This would take a fixed time, regardless of the size of the positive set, and could account for the increase in the zero-intercept. The second added stage might be the serial transfer of each item in the positive set into active memory, with a fixed average time per item transferred, estimated from the data to be about 50 msec. Since all items would be retransferred, whether the required response was positive or negative, the slopes of the functions for both responses would be increased by the same amount. The high-speed scan-

ning stage, which we already know to be exhaustive, would follow. The two added stages are not implausible and would account for the important features of the data. But this explanation—particularly the concept of “transferring a set of items into active memory”—needs to be made more precise and then tested.

11. An Explanation of Exhaustiveness

As mentioned in Sec. 3, an exhaustive search must involve more comparisons, on the average, than a search that terminates when a match occurs. The exhaustiveness of the high-speed scanning process therefore appears inefficient, and hence implausible. Why continue the comparison process beyond the point at which a match occurs? Figure 18 illustrates a system in which an exhaustive search would be more efficient than a self-terminating one for performance in

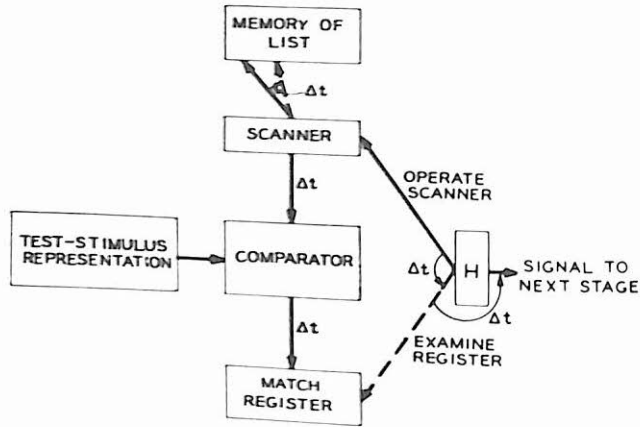


Figure 18

A system in which exhaustive scanning could be more efficient than self-terminating scanning. Some loci of possible time delays are represented by Δt s.

an item-recognition task. A representation of the test stimulus is placed in a comparator. When the scanner is being operated by the "central processor" or "homunculus," H, it delivers memory representations of the items in the list, one after another, to the comparator. If and when a match occurs a signal is delivered to the match register. The important feature of the system is that the homunculus can *either* operate the scanner *or* examine the register. It cannot engage in both of these functions at once, and switching between them takes time.

In this kind of system, if the switching time is long relative to the scanning rate, and if the list is sufficiently short, then an exhaustive search (in which the match register must be examined only once) is more efficient than a self-terminating one (where the register would have to be examined after each comparison). The surprisingly high speed of the scanning process may therefore be made possible by its exhaustiveness. But such a system might have at least one important limitation. After the search was com-

pleted, there might be no information available (without further reference to the memory of the list) as to the location in the list of the item that produced the match. The limitation would create no difficulty if the response required of the subject depended only on the presence or absence of an item in the list and not on its location, as in the item-recognition task. But the possibility that high-speed scanning does not yield location information does suggest an experiment to test this theory of exhaustiveness. Suppose we require a subject to give a response that *does* depend on where in the list a matching item is located. Then after each comparison, with information still available as to the location of the items just compared to the test stimulus (e.g., preserved by the position of the scanner in Fig. 18), it would have to be determined whether this item produced a match (by the homunculus switching from scanner to register). Scanning should then be slower than when only presence or absence has to be judged; it should also be self-terminating, since

further comparisons after a match had been detected would be superfluous. Such a process will be called *scanning to locate*.

12. Retrieval of Contextual Information by Scanning to Locate

In Fig. 19 is shown the paradigm of a *context-recall task*, one of the experiments devised to test these ideas (Exp. 6). On each trial the subject memorized a new random list of from three to seven different digits, presented visually one after another. The length of the list was varied at random from trial to trial. After a delay and a warning signal, a test item was presented, randomly selected from among all the digits in the list except the last. The test item, then, was always present in the list. The correct response was the spoken name of the item that followed the test item in the memorized list. The idea was that in order to make this response—that is, to recall an item defined by its contextual relation to the test item—the location of the test item in the

list might first have to be determined. As in the other experiments described, subjects were encouraged to respond as rapidly as possible, while attempting to maintain a low error rate.

Two aspects of the data are of particular interest: the relationship between mean RT and list length; and the relation, for a list of given length, between RT and the serial position of the test item in the list.

Data averaged over six subjects are shown in Fig. 20. Consider first Panel A. The bars show the percentage of wrong responses, which rises to 25% for lists of length 7. This is much higher than one would like, given an interest in error-free performance. The effect of list-length on mean RT is roughly linear, suggesting a scanning process. (Even closer approximations to linearity have been found in other similar experiments.) With a slope of 124 msec. per item, the fitted line is much steeper than the corresponding RT-function in the item-recognition task.

To interpret the slope, we have first to establish if the process is self-terminating, as expected. Evidence

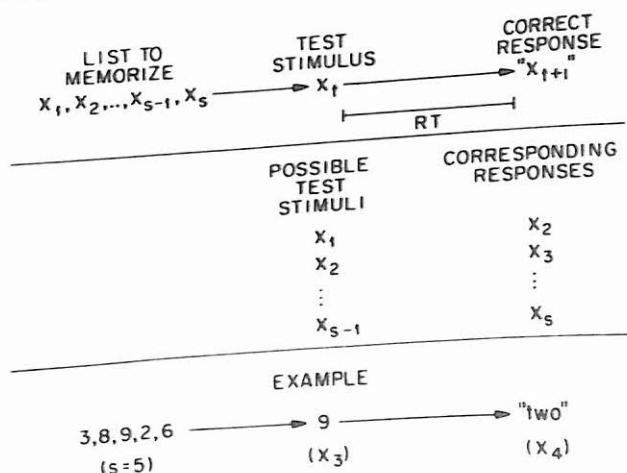


Figure 19

Paradigm of Exp. 6: Context-recall.

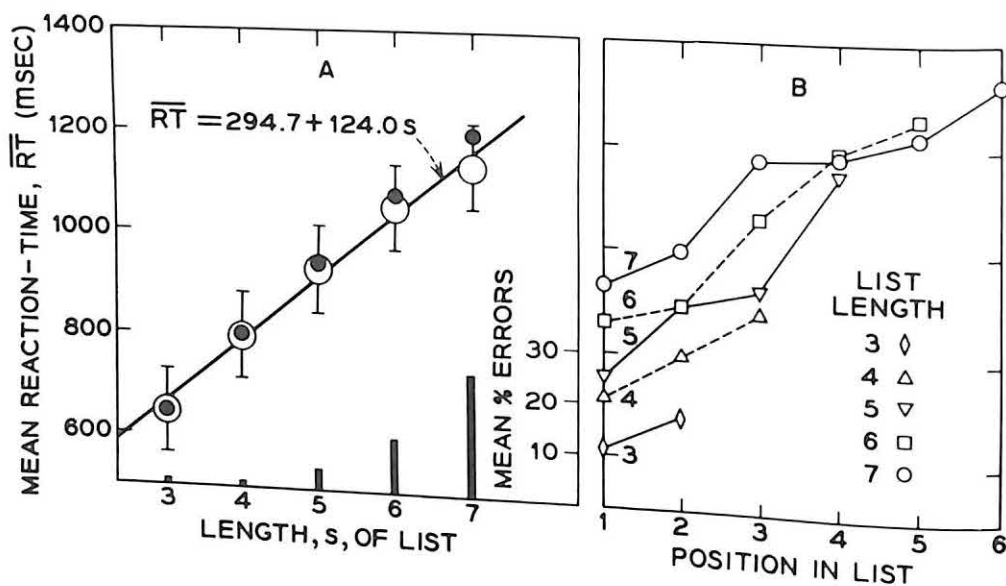


Figure 20

Results of Exp. 6: Context-recall. Averaged data from six subjects. A: Effect of list length on percent errors (bars), on mean latency of correct responses (open circles) with estimates of $\pm\sigma$ and line fitted by least squares, and on mean RT of all responses (filled circles). B: Relation between mean RT of correct responses and serial position of the test item in lists of five lengths.

on this point is provided by the average serial-position functions shown in Panel B. For each list length, mean RT is plotted as a function of the serial position of the test item in the memorized list. These functions are all increasing, suggesting a self-terminating process that tends to start at the beginning of the list and proceed in serial order.

Now we can interpret the slope of the function in Panel A, if we assume that list length influences only the scanning stage. (Evidence supporting this assumption of selective influence is presented below.) Since an average of about half the items in a list have to be scanned before a match occurs, the slope represents half of the time per item, and implies a scanning rate of about 250 msec. per item, or four items per second, in scanning to locate an item in a memorized list. Scanning to locate is therefore about seven

times as slow as the high-speed scanning process used to determine the presence of an item in a list. The slowness of the search, and the fact that it is self-terminating, lend support to the explanation (Sec. 11) of the exhaustiveness of the high-speed process. Scanning to locate seems to be fundamentally different from scanning for presence.¹²

¹² Alternative explanations of the dissimilarity of the two kinds of scanning are possible, of course. One interesting alternative (which existing data cannot reject) is that memory representations that can carry order information are different from those that need only carry item information, and that the observed differences in retrieval result from the fact that different kinds of memory representations are being scanned. However, for this alternative explanation to apply to Exp. 1 (in which subjects had to recognize an item and then recall the entire list in order), it must be possible for both kinds of memory representation to be maintained simultaneously.

As mentioned earlier (Sec. 3), if a self-terminating process started consistently at the beginning of a list and proceeded serially, the serial-position functions would be steep and superimposed, whereas if it started at a random point they would be flat and separated. The functions shown in Panel B lie between these extremes. This is partly because they represent averages of data from several subjects. Data from two subjects in Exp. 6 who represent almost pure cases, are shown in Fig. 21. The estimated scanning rates for these two subjects are almost the same, but their starting strategies appear to be radically different. Subject 1 seems to have started at a random point. This could occur if the presentation of the test item interrupted an ongoing cyclic rehearsal process, and scanning then began at the serial position where rehearsal happened to have stopped. Subject 4, on the other hand,

has the superimposed functions that would arise if he had started scanning consistently at the beginning of the list, perhaps by terminating his rehearsal before the test stimulus appeared. Data from other subjects range between these extremes, presumably because of mixed starting strategies.

One explanation of these results is the following: In order to recall a contextual item, the subject must first determine the test item's location in the memorized list. This is achieved by a slow, self-terminating process of scanning to locate, in which the items in memory are compared successively to the test item until a match occurs. Each nonmatching item that participates contributes to the RT a component time that depends neither on list length nor on the item's position in the list. This component time is occupied by switching to the item, comparing it to the test stimulus, and

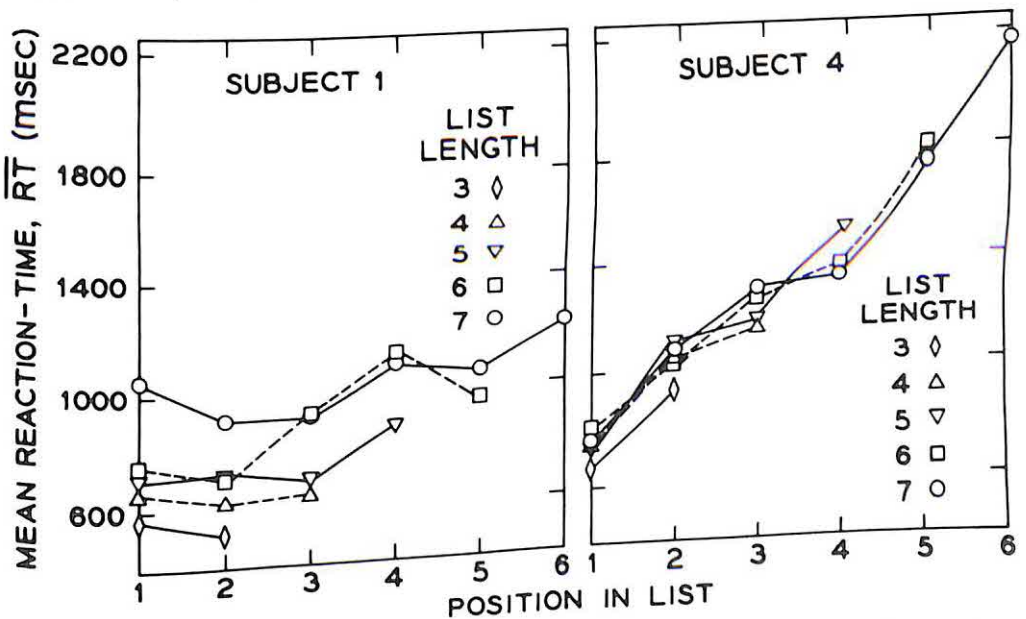


Figure 21

Individual data from Exp. 6: Context-recall. Contrasting sets of serial-position functions in lists of five lengths, one set relatively flat and separated, the other steep and, in general, superimposed.

determining that they have not matched. In the present context-recall task, the occurrence of a match is followed by a shift (e.g., a movement of the scanner in Fig. 18) from the item that matches the test item to the adjacent response item. For superimposed serial-position functions (as in Fig. 21) to be possible, we must assume that the duration of the shift operation (as well as other stages, such as stimulus-encoding and response-organization) is independent of the length of the list. Given this assumption, the slope of the RT-function is determined solely by the scanning rate.

The process of scanning to locate is a still more dramatic instance of having to hunt for information even when it is contained in a list that is being rehearsed. In some important sense one does not know what is in one's active memory, other than the single item to which attention is currently directed.¹³

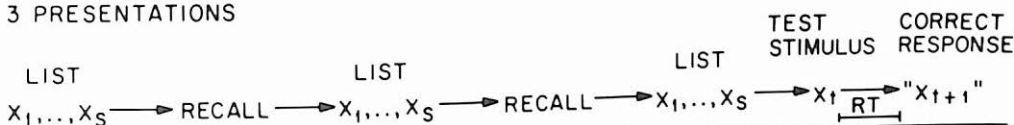
¹³ One traditional view is that the structure of a memorized list is a chain of overlapping associated pairs of items: the subject's task in a context-recall experiment is thought of as the performance of one of the associations in the chain, and the RT measure as an index of associative strength. At the least, this view must be modified to recognize the existence of a search for the representation of the test item in the list. This search is an instance of the obligatory process (usually ignored by association theorists) that locates and activates the memory trace of a stimulus before an associative response to that stimulus can be performed (Rock, 1962). Furthermore, in this experiment, not only does the locating process produce the dominant effect, but also there appears to be *no* influence of associative strength (Sec. 13). One might therefore question whether the traditional view is at all appropriate, at least for lists contained in active memory. It has been challenged from other directions also in recent years (e.g., Slamecka, 1967).

13. Independence of Learning and Retrieval from Active Memory

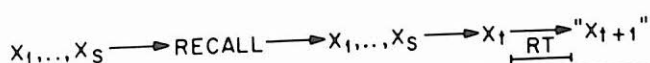
One problem with Exp. 6 is the high error rate, and its marked increase with list length (Fig. 20A). This makes the RT data somewhat suspect and violates the aim of studying error-free processes. Moreover, it raises the possibility that the level of learning of the list, which is clearly lower for longer lists, might be contributing to the increase of RT with list length. (For example, suppose that a list embodies a chain of associations and that the recall of a contextual item involves the performance of one of the associations. If the associations in a longer list are weaker, then at least one of the sources of the effect of list length on RT might be an increased associative latency.) In an experiment (Exp. 7) devised to look into these matters, the list was presented once, twice, or three times, as shown in Fig. 22, to vary how well it was learned. In the one-presentation condition, at the bottom of the figure, the list was presented, and there followed a test stimulus and response, just as in Exp. 6. Again, the list changed from trial to trial, and contained from 3 to 7 digits. In the two-presentation condition, each trial included an additional presentation of its list and an attempt to recall it. In the three-presentation condition there was still another presentation and recall of the list.

Results from six subjects are shown in Fig. 23. At the bottom, the percentage of errors in naming the succeeding digit is shown as a function of list length, for each condition. Added presentations reduced the error rate by a factor of three. At the top of the figure, mean RT is shown as a function of list length, for each of the

3 PRESENTATIONS



2 PRESENTATIONS



1 PRESENTATION

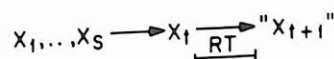


Figure 22

Conditions in Exp. 7: Effect of learning on context-recall.

three conditions. Despite the change in level of learning indicated by the error data, the pattern of RTs shows no systematic change with number of presentations.

This invariance indicates that dif-

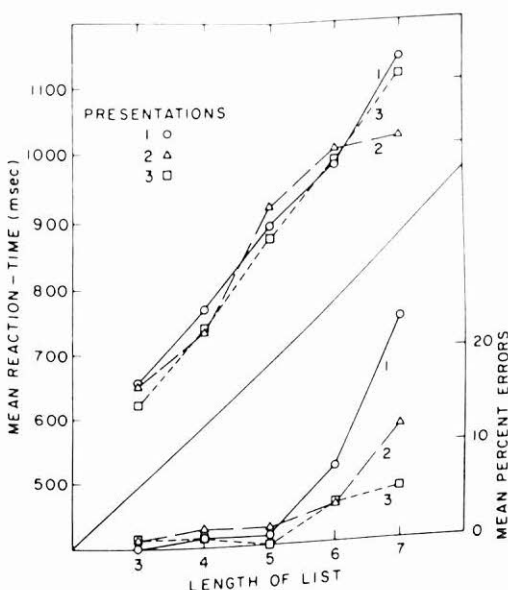


Figure 23

Results of Exp. 7: Effect of learning on context-recall. Averaged data from six subjects for one, two, and three presentations of the list. Bottom: mean percent errors in naming contextual item. Top: mean latency of correct responses.

ferences in level of learning that are associated with list length do not contribute to the influence of list length on mean RT, and further, that within the limits of the experiment, the rate of scanning to locate is independent of how well a list has been learned.¹⁴ The invariance with level of learning, which is similar to that of the high-speed scanning process over fixed-set and varied-set procedures, is consistent with the interpretation of the context-recall data presented in Sec. 11, and adds to the evidence that factors well known to influence learning may have no effect on active-memory functioning. Finally, given the invariance of the retrieval process, the strong influence of number of presentations on error rate suggests that the errors result primarily from faults in learning and retention, rather than in retrieval.

14 If a list could only be either perfectly learned or not learned at all, this conclusion would not be justified, since restricting the latencies analyzed to those of correct responses would entail the selection of lists that had been learned to the same degree (perfectly) in the three conditions. This objection does not apply here, mainly because correct responses in conjunction with *partially*-learned lists were frequent.

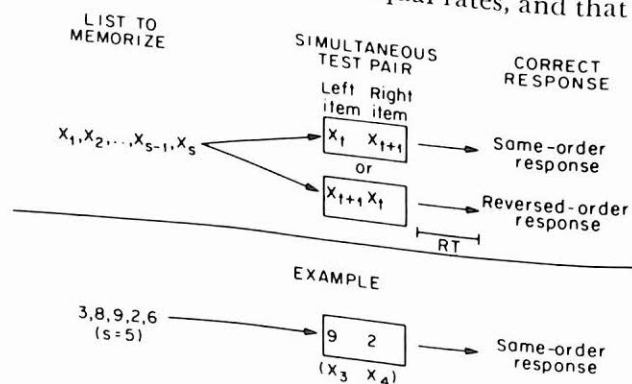
14. Recall Versus Recognition of Contextual Information

In explaining the difference between findings from the item-recognition and context-recall tasks (Exps. 1-5 *versus* Exps. 6-7) I have emphasized that in one case the response depends merely on the presence of an item in the list, and in the other case on its exact location. For an explanation in these terms to be valid, however, certain other differences between the tasks must be shown to be unimportant: one task involves recall, and the other recognition; one requires that for production of the response a memory representation be converted into a particular form—its name—and the other does not; and whereas the number of response alternatives in one task grows with list length, the other always requires a binary choice.

The last experiment to be described (Exp. 8) was designed to evaluate the importance of these factors and to begin to examine further the generality of the process of scanning to locate. A recognition procedure was used to study the retrieval of contextual information; the resulting *context-recognition task* is shown in Fig. 24. On each trial the subject attempted to

memorize a list of from 3 to 6 different digits, presented visually, one after another. To increase accuracy, the list was actually presented twice, with a recall attempt after the first presentation. The test stimulus was a pair of simultaneously presented digits that had appeared successively somewhere in the list. The subject's task was to decide whether the left-to-right order of the pair was the same as its temporal order in the list, or reversed. He made his response by pulling one of two levers (as in Exps. 1-5).

This experiment seemed to be somewhat risky, since there appeared to be a variety of strategies open to the subject. One possibility was that before its order could be tested, the pair might have to be located in the list by means of a scanning process. This process would be revealed by the relation between RT and the length of the list. Suppose that the test pair is located in the list by scanning for the location of one of its members, according to the self-terminating process described in Sec. 11. One would then expect that in the context-recognition task mean RT for both same-order and reversed-order responses would increase linearly with list length, and at equal rates, and that the



Paradigm of Exp. 8: Context-recognition.

Figure 24

rate of increase would be the same as in the context-recall task.

The same six subjects who performed the context-recall task of Exp. 7 also served (in a balanced order) in the recognition task. RT-functions for both responses (Fig. 25)

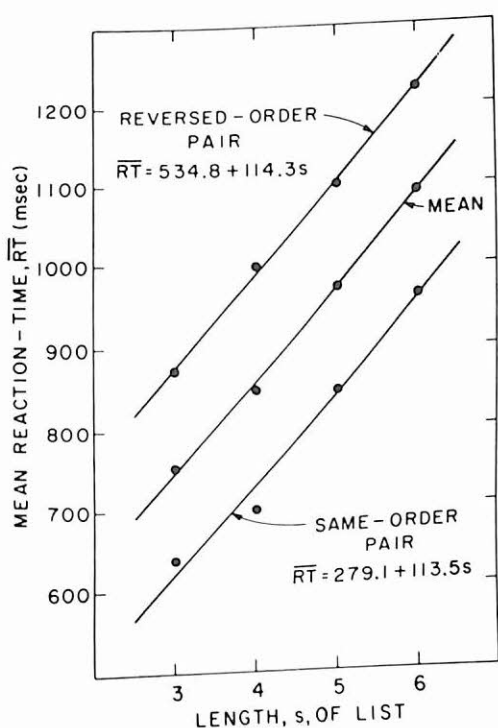


Figure 25

Results of Exp. 8: Context-recognition. Averaged data from six subjects. Mean latencies of correct same-order and reversed-order responses, and their mean, with lines fitted by least squares.

are linear, supporting the notion that in this task, also, performance involves a scanning process. For both responses the slope of the fitted line is 114 msec. per item.¹⁵ The equality

¹⁵ Although equal in slope, the RT-functions for the two kinds of response differ by about 250 msec. in intercept. The several ways in which one might account for this difference are not discussed in this paper.

of slopes is consistent with the idea that both responses depend on first locating one of the members of the pair in the list. That this is accomplished by means of a self-terminating process is suggested by the serial-position data: for all subjects, and for both responses, mean RT increased with the serial position of the pair in the list. Averaged serial-position data are shown in Fig. 26.

RT-functions from the context-recognition and context-recall tasks (Exps. 7 and 8) are compared in Fig. 27. The fitted lines are parallel, supporting the idea that the same search process (scanning to locate) underlies performance in both tasks. Also shown, for reference, is the RT-func-

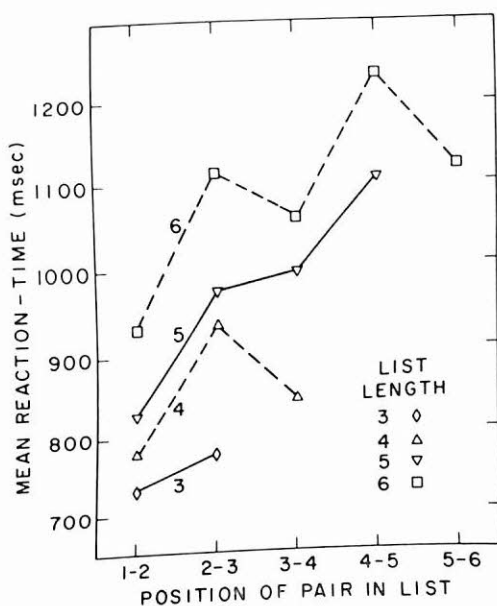


Figure 26

Further results of Exp. 8: Context-recognition. Relation between mean RT of correct responses and serial position of the test pair in lists of four lengths. Data were averaged over six subjects and over same-order and reversed-order responses.

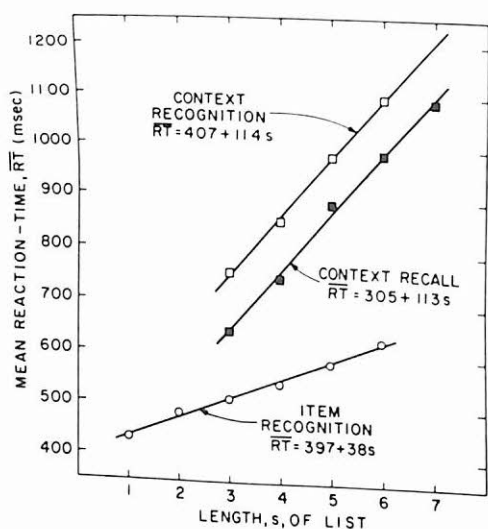


Figure 27

Comparison of results from context-recognition, context-recall, and item-recognition tasks. Top: Mean RTs from Exp. 8 (Fig. 24), averaged over same-order and reversed-order responses. Middle: Mean RTs averaged over the three conditions in the context-recall task of Exp. 7 (Fig. 22), which used the same subjects as Exp. 8. Bottom: Mean RTs from Exp. 1 (Fig. 4).

tion from the item-recognition task of Exp. 1.

These parallel lines provide another striking instance of additive effects on mean RT. Here the additive factors are task (context-recognition *versus* recall) and list length; the absence of interaction indicates that these factors influence processing stages selectively, and helps to justify our interpretations of the data. Apparently, the change from recall to recognition does not influence the scanning stage, and, as assumed in Sec. 12, changes in list length do not influence perceptual and response stages.

One final substantive point about these results concerns their implications for the recognition-recall distinction. It is tempting to think that recognition involves less search, in some sense, than recall. These data reveal at least one search process that is as evident in a recognition task (Exp. 8) as in a recall task (Exp. 7).

15. Concluding Remarks

I have described several experiments whose analysis is based on a modern version of Donders' subtraction method. These experiments have led to the discovery of two kinds of scanning process, with rather different structures and rates. Which of these processes is used in retrieving information from a given memorized list depends on the kind of information sought.

Among other substantive implications of the experiments are: (1) Apparently one *must* scan a list serially to retrieve information from it, even when it is contained in active memory. There is no evidence in any of these data that one can "think about" more than one thing at a time, and thereby simultaneously compare a set of memorized items to a test item. (2) On the other hand, even a well-learned list can be made more readily available by being maintained in active memory. (3) Despite the possibility that retention may depend on a rehearsal process involving covert speech, visual representations are used for comparison to representations of visual stimuli. (4) The same search process can be involved in both recall and recognition tasks.

In the course of describing these experiments, I have indicated how a theory of processing stages, together

with an assumption of selective influence, lead to methods of experimentation and data analysis that may overcome certain weaknesses in earlier applications of the subtraction method. These weaknesses resulted from the use, without test, of a strong assumption of pure insertion. The basic ideas behind the modern approach are, first, that (1) if separate stages between stimulus and response have been correctly identified, then for each of these stages it may be easier to find a factor that *influences* it without altering other stages than to find one that *inserts* it without altering other stages; and (2) these factors would then have additive effects on mean RT. The discovery of several sets of such additive factors was critical in the interpretation of the experiments described.

References

- ATKINSON, R. C. & SHIFFRIN, R. M., Human memory: a proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The Psychology of learning and motivation: Advances in research and theory*. Vol. 2. New York: Academic Press, 1968. Pp. 89-195.
- BADDELEY, A. D., The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quart. J. exp. Psychol.*, 1966, 18, 302-309.
- BROADBENT, D. E., *Perception and communication*. London: Pergamon Press, 1958.
- CATTELL, J. McK., The perception time. *Mind*, 1886, 11, 377-392. Reprinted in *James McKeen Cattell, Man of science*. Lancaster, Pa.: The Science Press, 1947. Pp. 64-79.
- CHASE, W. G. & CALFEE, R. C., Modality and similarity effects in short-term recognition memory. *J. exp. Psychol.*, 1969, 81, 510-14.
- COHEN, R. L. & JOHANSSON, B. S., Some relevant factors in the transfer of material from short-term to long-term memory. *Quart. J. exp. Psychol.*, 1967, 19, 300-308.
- CONRAD, R., Acoustic confusions in immediate memory. *Brit. J. Psychol.*, 1964, 55, 75-84.
- CORBALLIS, M. C., Serial order in recognition and recall. *J. exp. Psychol.*, 1967, 74, 99-105.
- CROWDER, R. G., Short-term memory for words with a perceptual-motor interpolated activity. *J. verb. Learn. verb. Behav.*, 1967, 6, 753-761.
- DONDERS, F. C., Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868-1869, Tweede reeks, II, 92-120. Transl. by W. G. Koster in W. G. Koster (Ed.), *Attention and performance II*. *Acta Psychol.*, 1969, 30, 412-431.
- EGETH, H. E., Parallel versus serial processes in multidimensional stimulus discrimination. *Percept. & Psychophys.*, 1966, 1, 245-252.
- GLANZER, M. & CUNITZ, A. R., Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.*, 1966, 5, 351-360.
- HOCHBERG, J., In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart & Winston, 1968. Pp. 309-331.
- HOHLE, R. H., Component process latencies in reaction times of children and adults. In L. P. Lipsett and C. C. Spiker (Eds.), *Advances in child development and behavior*. Vol. 3. New York: Academic Press, 1967. Pp. 225-261.

- JASTROW, J., *The time-relations of mental phenomena. Fact & theory papers No. VI.* New York: N.D.C. Hodges, 1890.
- JOHNSON, D. M., *The psychology of thought and judgment.* New York: Harper, 1955.
- KÜLPE, O., *Outlines of psychology.* New York: Macmillan, 1895.
- LANDAUER, T. K., Rate of implicit speech. *Percept. Mot. Skills*, 1962, 15, 646.
- LANDAUER, T. K. & FREEDMAN, J. L., Information retrieval from long-term memory: category size and recognition time. *J. verb. Learn. verb. Behav.*, 1968, 7, 291-295.
- MCGILL, W. J. & GIBBON, J., The general gamma distribution and reaction times. *J. math. Psychol.*, 1965, 2, 1-18.
- McMAHON, L. E., Grammatical analysis as part of understanding a sentence. Unpublished doctoral dissertation, Harvard University, 1963.
- MILLER, G. A., *Psychology, the science of mental life.* New York: Harper & Row, 1962.
- MORIN, R. E., DeROSA, D. V. & STULTZ, V., Recognition memory and reaction time. In A. F. Sanders (Ed.), *Attention and performance. Acta Psychol.*, 1967, 27, 298-305.
- MORIN, R. E., DeROSA, D. V. & ULM, R., Short-term recognition memory for spatially isolated items. *Psychon. Sci.*, 1967, 9, 617-618.
- NEISSER, U., *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.
- NEWELL, A. & SIMON, H. A., Computers in psychology. In R. D. Luce, R. R. Bush and E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963. Pp. 361-428.
- NICKERSON, R. S., Categorization time with categories defined by disjunctions and conjunctions of stimulus attributes. *J. exp. Psychol.*, 1967, 73, 211-219.
- POSNER, M. I., Short-term memory systems in human information processing. In A. F. Sanders (Ed.), *Attention and performance. Acta Psychol.*, 1967, 27, 267-284.
- POSNER, M. I., BOIES, S. J., EICHELMAN, W. H., & TAYLOR, R. L., Retention of visual and name codes of single letters. *J. exp. Psychol. Monogr.*, 1969, 79, No. 1, Part 2, 1-16.
- POSNER, M. I. & MITCHELL, R. F., Chronometric analysis of classification. *Psychol. Rev.*, 1967, 74, 392-409.
- POSNER, M. I. & ROSSMAN, E., Effect of size and location of informational transforms upon short-term retention. *J. exp. Psychol.*, 1965, 70, 496-505.
- RESTLE, F. & DAVIS, J. H., Success and speed of problem solving by individuals and groups. *Psychol. Rev.*, 1962, 69, 520-536.
- ROCK, I., A neglected aspect of the problem of recall: The Höfdding function. In J. M. Scher (Ed.), *Theories of the mind.* New York: Free Press, 1962. Pp. 645-659.
- SANDERS, A. F., Rehearsal and recall in immediate memory. *Ergonomics*, 1961, 4, 25-34.
- SLAMECKA, N. J., Serial learning and order information. *J. exp. Psychol.*, 1967, 74, 62-66.
- SMITH, E. E., Effects of familiarity on stimulus recognition and categorization. *J. exp. Psychol.*, 1967, 74, 324-332.
- SMITH, E. E., Choice reaction time: an analysis of the major theoretical positions. *Psychol. Bull.*, 1968, 69, 77-110.
- SPERLING, G., The information available in brief visual presentations. *Psychol. Monogr.*, 1960, 75 (11, Whole No. 498).
- SPERLING, G., A model for visual memory tasks. *Hum. Factors*, 1963, 5, 19-31.
- STERNBERG, S., Stochastic learning theory. In R. D. Luce, R. R. Bush and E.

- Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 2. New York: Wiley, 1963. Pp. 1-120.
- STERNBERG, S., Estimating the distribution of additive reaction-time components. Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ont., October 1964.
- STERNBERG, S., High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- STERNBERG, S., Retrieval of contextual information from memory. *Psychon. Sci.*, 1967, 8, 55-56. (a)
- STERNBERG, S., Two operations in character-recognition: some evidence from reaction-time measurements. *Percept. & Psychophys.*, 1967, 2, 45-53. (b)
- STERNBERG, S., The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. *Acta Psychol.*, 1969, 30, 276-315.
- SUPPES, P. & GROEN, G., Some counting models for first-grade performance data on simple addition facts. Technical Report 90, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1966.
- TAYLOR, D. H., Latency components in two-choice responding. *J. exp. Psychol.*, 1966, 72, 481-488.
- WAUGH, N. C. & NORMAN, D. A., Primary memory. *Psychol. Rev.*, 1965, 72, 89-104.
- WICKELGREN, W. A., Auditory or articulatory coding in verbal short-term memory. *Psychol. Rev.*, 1969, 76, 232-235.
- WOODWORTH, R. S., *Experimental psychology*. New York: Holt, 1938.
- WUNDT, W., *Grundzüge der physiologischen Psychologie*, Vol. II, 2nd ed. Leipzig: Engelmann, 1880.
- YONAS, A., The acquisition of information-processing strategies in a time-dependent task. Unpublished doctoral dissertation, Cornell University, 1969.



NOTES

NOTES

Form No. 3.

PSY, RES.L-1

**Bureau of Educational & Psychological
Research Library.**

The book is to be returned within
the date stamped last.

.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		
.....		

Form No. 4

BOOK CARD

Coll. No..... Accn. No.....

Author.....

Title.....

Date.	Issued to	Returned on
.....		

150
GAZ

good
reading
in
psychology

edited by **MICHAEL S. GAZZANIGA** and **ELIJAH P. LOVEJOY**

The essay "On War and Peace in Animals and Man" by N. Tinbergen illustrates two important criteria for inclusion in this new collection: it is integrative and it is interesting. Like other essays in the book, it takes assorted data and conclusions from a wide-range of studies and integrates this information into a comprehensive and smoothly-reading essay.

The editors narrow the coverage of this book to a small number of especially exciting areas to afford the reader an in-depth view of each area. They also seek to present "the very best that the field has to offer" by searching for outstanding works of great thinkers and writers.

Outstanding Features:

- In-depth studies of a few areas, rather than a shallow attempt "to cover"
- Essays by men who are leaders in their respective fields. Newcomers to the field are more likely to respect the opinions of men whose names they recognize
- Topics of current interest to readers
- Introductions to all the essays...These focus the reader's attention on special points of interest and on interrelationships among papers. They also help the reader understand the more technical papers.
- Wide-ranging and integrative essays

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey

13-360503-5